Genetic Evaluation Models and Strategies for Potato Variety Selection

A thesis

submitted in partial fulfilment

of the requirements for the Degree

of

Doctor of Philosophy

in the

University of Canterbury

Mark F. Paget

University of Canterbury

2014

Abstract

A series of studies are presented on the genetic evaluation of cultivated potato (*Solanum tuberosum* L.) to improve the accuracy and efficiency of selection at various stages of a breeding programme. The central theme was the use of correlated data, such as relationship information and spatial and across-trial correlations, within a linear mixed modelling framework to enhance the evaluation of candidate genotypes and to improve the genetic response to selection. Analyses focused on several social and economically-important traits for the enhancement of the nutritional value, disease resistance and yield of potato tubers.

At the formative stages of a breeding scheme, devising a breeding strategy requires an improved understanding of the genetic control of target traits for selection. To guide a strategy that aims to enhance the micronutrient content of potato tubers (biofortification), univariate and multivariate Bayesian models were developed to estimate genetic parameters for micronutrient tuber content from a breeding population generated from crosses between Andean landrace cultivars. The importance of the additive genetic components and extent of the narrow-sense heritability estimates indicated that genotypic 'individual' recurrent selection based on empirical breeding values rather than family-based selection is likely to be the most effective strategy in this breeding population. The magnitude of genetic correlations also indicated that simultaneous increases in important tuber minerals, iron and zinc, could be achieved.

Optimising selection efficiency is an important ambition of plant breeding programmes. Reducing the level of candidate replication in field trials may, under certain circumstances, contribute to this aim. Empirical field data and computer simulations inferred that improved rates of genetic gain with p-rep (partially replicated) testing could be obtained compared with testing in fully replicated trials at the early selection stages, particularly when testing over two locations. P-rep testing was able to increase the intensity of selection and the distribution of candidate entries across locations to account for G×E effects was possible at an earlier stage than is currently practised. On the basis of these results, it was recommended that the full replication of trials (at the first opportunity, when enough planting material is available) at a single location in the early stages of selection should be replaced with the partial replication of selection candidates that are distributed over two locations.

Genetic evaluation aims to identify genotypes with high empirical breeding values (EBVs) for selection as parents. Using mixed models, spatial parameters to target greater

control of localised field heterogeneity were estimated and variance models to account for across-trial genetic heterogeneity were tested for the evaluation of soil-borne powdery scab disease and tuber yield traits at the early stages of a selection programme. When spatial effects improved model fit, spatial correlations for rows and columns were mostly small for powdery scab, and often small and negative for marketable and total tuber yield suggesting the presence of interplot competition in some years for tuber yield traits. For the evaluation of powdery scab, genetic variance structures were tested using data from 12 years of long-term potato breeding METs (multi-environment trials). A simple homogeneous correlation model for the genetic effects was preferred over a more complex factor analytic (FA) model. Similarly, for the MET evaluation of tuber yield at the early stages, there was little benefit in using more complex FA models, with simple correlation structures generally the most favourable models fitted. The use of less complex models will be more straightforward for routine implementation of potato genetic evaluations in breeding programmes.

Evaluations for (marketable) tuber yield were extended to multi-location MET data to characterise both genotypes and environments, allowing a re-evaluation of New Zealand MET selection strategies aimed at broad adaptation. Using a factor analytic mixed model, results indicated that the programme's two main trial locations in the North and the South Islands optimised differentiation between genotypes in terms of G×E effects. There was reasonable performance stability of genotypes across test locations and evidence was presented for some, but limited, genetic progress of cultivars and advanced clonal selections for tuber marketable yield in New Zealand over recent years.

The models and selection strategies investigated and developed in this thesis will allow an improved and more systematic application of genetic evaluations in potato selection schemes. This will provide the basis for well informed decisions to be made on selection candidates for the genetic improvement of potato in breeding programmes.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 2. Genetic evaluation of micronutrient traits in diploid potato from a base population of Andean landrace cultivars. M. Paget, W. Amoros, E. Salas, R. Eyzaguirre, P. Alspach, L. Apiolaza, A. Noble, M. Bonierbale. Crop Science (in press). doi: 10.2135/cropsci2013.12.0809

Please detail the nature and extent (%) of contribution by the candidate:

90% data analysis. 90% writing of the manuscript.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Luis Apiolaza Signature:  Date: 27th March 2014

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Chapter 4. Genetic variance models for the evaluation of resistance to powdery scab (Spongospora subterranea f. sp. subterranea) from long-term potato breeding trials. M. Paget, P. Alspach, R. Genet, L. Apiolaza. Euphytica 197(3): 369-385 (2014). doi: 10.1007/s10681-014-1073-9

Please detail the nature and extent (%) of contribution by the candidate:

90% data analysis. 90% writing of the manuscript.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all

The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: Luis Apiolaza Signature:  Date: 27th March 2014

Acknowledgements

Thesis work is a prolonged endeavour and there are numerous people that have provided me with the support, instruction and inspiration to help me complete this venture. These people deserve special mention.

I would therefore like to take this opportunity to express my gratitude to my senior supervisor Luis Apiolaza, whose support and guidance has always been helpful and willingly offered. Likewise, Peter Alspach, colleague and co-supervisor, deserves special thanks as an enthusiastic source of advice, insight and encouragement throughout. My appreciation must also go to Russell Genet whose knowledge of potato breeding has been acquired over many years. This was a great resource when I needed a better understanding of the practical implications of potato breeding research.

Merideth Bonierbale, Walter Amoros and Elisa Salas, at the International Potato Center (CIP) in Lima, Peru, were very generous in their hospitality and also provided a wealth of knowledge with regard to potato breeding and genetics. I am sincerely grateful to them for allowing me access to and exploration of their data.

Jeanette Allen of the School of Forestry at the University of Canterbury provided thorough guidance and clarification on all administrative requirements, for which I am grateful. I would also like to extend my gratitude to a number of staff members who mostly reside at The New Zealand Institute of Plant & Food Research Limited: to John Anderson and Fred Braam for their help in organising and managing field trials and collecting data at Pukekohe and Lincoln; to Russell and John once again for providing their many years of New Zealand trial data; to Alasdair Noble and Raul Eyzaguirre (CIP) for their advice on using Bayesian methods and Ruth Butler for her guidance on field trial design in general and the use of trial design software in particular; to Donna Gibson and Robert Lamberts for photographs; to Naomi Shaw for locating elusive articles and books, and for help with formatting the bibliography; and to Ron Beatson, Richard Falloon, Anne Gunson, Steve Lewthwaite, Carolyn Lister, Sathiyamoorthy ‘Mei’ Meiyalaghan, Gail Timmerman-Vaughan and Richard Volz for their miscellaneous advice on various aspects of this study.

I gratefully acknowledge the generous financial support from the Potatoes NZ Charitable Trust. My appreciation also goes to The New Zealand Institute of Plant & Food Research Limited and, in particular, Bill Griffin and Tony Conner for providing me with the opportunity and privilege to study towards this degree.

To Frances, James, Heather and Samuel

Declaration

I declare that the material presented in this thesis is my own work, with the following exceptions:

i. Alasdair Noble provided the R code for analysis using the package 'rjags' in Chapter 2, (p. 42).

ii. Peter Alspach helped to develop the method and R code for the bootstrap resampling procedure outlined in Chapter 3, (p. 52).

iii. Luis Apiolaza and Peter Alspach provided miscellaneous pieces of R code that were used throughout the thesis.

iv. Trial data was collected by staff at the International Potato Centre (CIP) in Chapter 2 and The New Zealand Institute for Plant & Food Research in Chapters 3, 4, 5 and 6.

Material presented in Chapter 2 has previously been published (online, to date) in Paget, M. et al. (2014) Genetic evaluation of micronutrient traits in diploid potato from a base population of Andean landrace cultivars. Crop Science (*in press*). doi: 10.2135/cropsci2013.12.0809

Material presented in Chapter 4 has previously been published in Paget, M. et al. (2014) Genetic variance models for the evaluation of resistance to powdery scab (*Spongospora subterranea* f. sp. *subterranea*) from long-term potato breeding trials. Euphytica 197(3): 369-385. doi: 10.1007/s10681-014-1073-9

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike Information Criterion
AMMI	additive main effects and multiplicative interaction (model)
ANOVA	analysis of variance
BLUP	Best Linear Unbiased Prediction
BLUE	Best Linear Unbiased Estimation
CGIAR	Consultative Group on International Agricultural Research
CHI	commercial harvest index (percent marketable yield)
CIS	cold-induced sweetening
CIP	International Potato Centre (Centro Internacional de la Papa)
DIC	Deviance Information Criterion
EBV	empirical (or estimated) breeding value
FA	factor analytic
FAO	Food and Agricultural Organization of the United Nations
GCA	general combining ability
GS	genomic selection
GWAS	genome-wide association studies
G×E	genotype-by-environment
IB	incomplete block (design)
LRT	Likelihood Ratio Test
MAS	marker-assisted selection
MCMC	Markov chain Monte Carlo
MET	multi-environment trials
ML	maximum likelihood
MME	mixed model equations
MTY	marketable tuber yield
MVN	multivariate normal
NCD	North Carolina Design
PCN	potato cyst nematode
PEV	prediction error variance
PFR	The New Zealand Institute for Plant & Food Research
PVR	Plant Variety Rights
PVX/PVY/PLRV	potato virus X / potato virus Y / potato leaf roll virus
QTL	quantitative trait loci
RCBD	randomised complete block design
REML	Residual (or Restricted) Maximum Likelihood
SCA	specific combining ability
TTY	total tuber yield

To clarify terms used throughout the thesis, variety is a generic term that can refer to a genotype, entry, candidate, clone, breeding line or cultivar; genotype, entry, candidate, clone and breeding line are synonymous; and a cultivar is either a variety that has been commercially released with Plant Variety Rights (PVR) or is a native landrace variety.

The statistical convention of representing ‘estimates’ with a ‘hat’ has been generally ignored for the sake of simplicity.

1 General Introduction

1.1 The potato: a world food staple

The cultivated potato is a major food source in many regions and, along with the grain staples wheat, rice and maize, is considered one of the top four staple food crop of global importance. In 2012, global production exceeded 368 million tonnes from just over an estimated 19 million hectares (FAOSTAT 2013). The plant is native to South America and has been a part of the Andean diet for thousands of years. After its introduction into Europe in the 16th century, it quickly became an important diet mainstay, with many historians and writers claiming it to have helped to drive the industrial revolution, European migration and colonial expansion (e.g. Hobhouse 1987; McNeill 1999; Zuckerman 1999). In the past 50 years, its status in many regions of Africa and Asia has been increasing (Hijmans 2001) and in 2005, the FAO (Food and Agriculture Organization of the United Nations) reported that potato production in developing nations exceeded that of the developed world for the first time; output in Asia, Africa and Latin America rose from less than 30 million tonnes in the 1960s to 165 million tonnes in 2007 (FAO 2010), illustrating its significance as a global staple.

Potatoes are a valuable source of nutrients, producing more energy and protein per hectare of land than any grain crop (Bamberg and del Rio 2005). They also contain a number of important minerals, vitamins and antioxidant phytochemicals (Andre et al. 2007; Brown 2008) including vitamin C, vitamins B_3 and B_6, potassium, phosphorus, magnesium, carotenoids and polyphenols. Their widespread cultivation has also been attributed to their ability to grow in environments that vary considerably in terms of latitude, altitude, daylength and temperature (Hijmans 2001; Veilleux and De Jong 2007). Potatoes are consumed either as a fresh product or processed e.g. as French fries or crisps. They can also be a source of biofuel as an alternative to fossil fuels and starch for the production of biodegradable composites, so they are very versatile.

1.2 The origin of modern cultivars

The germplasm base of potato is comprised of primitive indigenous (cultivated) landraces and wild *Solanum* species (Ovchinnikova et al. 2011). Advances in molecular techniques have provided greater insight into the taxonomic relationships of potato and currently there

are estimated to be 100 wild species and four cultivated species (Spooner et al. 2007). *Solanum tuberosum* L. is the major cultivated species of potato, which was domesticated about 7,000 to 10,000 years ago in the Andes of Peru (Hawkes 1990; Spooner et al. 2005). It is a highly heterozygous outcrossing species which is asexually propagated, via tubers, for food production and germplasm maintenance. Sexual propagation and the production of 'true' seed allow breeders to generate genetic variation, and as a clonal crop, there are opportunities to exploit both additive and non-additive variation (Mackay 2007). Several native taxonomic 'cultivar Groups' that are found in cultivated potato populations grown in the South American Andes exist within *S. tuberosum*, such as Groups Phureja, Stenotomum and Andigena. These vary in ploidy levels (diploid to pentaploid) and show a high degree of diversity in tuber size, shape, skin and flesh colour, storage ability and cooking quality (Andre et al. 2007). Andean landraces are important as sources of food and income to the communities living in this region. Many commercial cultivars widely grown globally are tetraploid (2n=4x=42) of Group Tuberosum, which have been adapted to form tubers in long-day conditions (i.e. in the main growing season of the higher and lower latitudes of the northern and southern hemispheres). Most modern cultivars are considered to be based on Chilean landraces (that were adapted to the long days of the lower southern latitudes) after their introduction into Europe in the early 19th century, which displaced the Andean potato that had predominated since their arrival in Europe in the 16th century (Ames and Spooner 2008). After European introduction, the potato was distributed to the far reaches of the globe and would have been one of the earliest food crops to become established in New Zealand after early European settlement, having first been introduced to the native Maori during the initial expeditions of British and French explorers (Reader 2011). The native Maori people were quick to appreciate the potential of this new exotic crop, not least because it could be successfully cultivated in the colder south (McNeill 1999; Reader 2011). This was in contrast to the hitherto established staple food, sweet potato (*Ipomoea batatas*), locally known as kumara, which had arrived with the early Pacific settlers and is a crop better adapted to the warmer climate found in the north of the islands.

As well as indigenous and primitive landraces, there are a number of wild species of the *Solanum* genus (section Petota) of various ploidy levels that can be used (with varying degrees of difficulty) as new sources of genetic diversity, for a range of economically important traits, to develop new *S. tuberosum* potato varieties (Hawkes 1990; Bradshaw and Ramsay 2005). Modern cultivars of *S. tuberosum* have been subjected to interspecific hybridisations and intensive breeding over the course of the 20th century. It is reported that

these have been developed from *S. tuberosum* landraces and 15 wild species, indicating that only a small proportion of wild species (~10%) has been used in modern potato breeding (Ross 1986; Plaisted and Hoopes 1989; Ovchinnikova et al. 2011). This is largely attributed to the difficulties of interspecific hybridisation and also the undesirable characteristics that are introduced from the donor alongside the target trait or traits, which may take many generations of further backcrossing with the recipient species to eliminate.

1.3 Breeding objectives and selection criteria

As a technological development, genetic improvement has been described by Groen (2003) as the saving of production factors that achieve a market price or an opportunity cost which can then be used, for example, to increase food security or food quality, increase well-being or improve system sustainability. The overall goal of a genetic improvement programme is defined by the breeding objective, which is the quantitative description of the relative or absolute benefits of improvements (i.e. the value given to saved production factors) in all genetic traits of interest (Amer 1995). This might be, for example, to maximise profit, to maximise economic efficiency, or to minimise risk. It is recognised that the beneficiary will depend upon the perspective from which the objective is quantified, e.g. the individual breeder/seed supplier, farmer, factory, industry or society as a whole. It is generally considered ideal to include all traits via the selection criteria that will contribute (directly or indirectly) to the objective. In practice, a compromise has to be met so traits are restricted to those that have potential for genetic change and those which are cost-effective to measure within a finite set of resources; '*... a breeding programme is ultimately dominated by the biological facts and possibilities'* (Simmonds 1979). Economic values (or weights) are derived to weight each trait appropriately and are then combined with their estimated breeding values in a selection index. Selection indices are considered as an optimal selection method, under certain assumptions, for the simultaneous improvement of multiple traits in a breeding objective (Falconer and Mackay 1996). From the literature, it seems apparent that plant breeders typically use intuitive selection methods which have an implicit economic objective, such as desired gains indices and independent culling. Explicit economic criteria in plant breeding are not widely reported, possibly for commercial reasons or simply that they are poorly developed, and rarely defined in the breeding objective (Simmonds and Walker 1986; Sölkner et al. 2008), with the possible exception of sugarcane (Simmonds 1979; Simmonds and Walker 1986; Deren et al. 1992).

In a potato improvement programme, the traits to consider in a breeding objective will ultimately vary, depending on the environments in which the crop is to be grown and the intended end-use. Potato production in New Zealand is concentrated in the Pukekohe, Waikato, Hawke's Bay, Manawatu and Canterbury regions, which are generally considered as having a temperate climate with relatively small variations in winter and summer temperatures (NIWA 2013). Presently, potato production covers approximately 10,700 hectares and grown for the table (fresh production) (~3,500 hectares), for processing into manufactured potato products (~5,900 hectares), or as seed potatoes (~1,200 hectares) (Potatoes NZ 2013). For potato in general, traits targeted in a breeding programme will include yield and tuber conformation, resistances to various biotic and abiotic stresses, cooking quality and storage ability. The overall aim is therefore to enhance productivity whilst meeting various specifications in terms of agronomic and quality characteristics for a particular set of environments and end-users. In New Zealand, there is a need for self-sufficiency in table (fresh) potato production because of the cost and biosecurity issues of importing potato tubers. There are also valuable export markets for processed potato products. The potato breeding programme, based at The New Zealand Institute for Plant & Food Research Limited (PFR), has developed cultivars for both the table and processing sectors in New Zealand and Australia and has clonal varieties currently in field trials in India, North Africa, the USA and Europe. Traits that are currently targeted in selection are highlighted in Table 1-1 and, in general, illustrate a typical perspective for many specialised production systems in Europe, North America and Australasia.

Tuber yield as an ongoing target for selection

Significant yield gains in staple crops for many global regions over the past century have been underpinned by technological developments in both genetics and agronomy (Kang 2002a). In New Zealand, for example, records show that potato tuber yield has increased from an estimate of 23 tonnes per hectare in 1961, based on a total production area of 11100 hectares, to approximately 47 tonnes per hectare in 2012 produced from 11500 hectares (FAOSTAT 2013). Such gains are typical of countries with specialised production systems, for example, those found in North America and Northern Europe where yield per hectare was reported to be over double the world average between 2001 and 2005 (Veilleux and De Jong 2007). Yield gains, however, have not necessarily been restricted to these regions, as Walker et al. (2003) reported a relative tuber yield increase approaching 40% in developing countries between 1980 and 2000. The factors that have influenced potato yield in New Zealand over

the past decades are difficult to ascertain and have not been closely examined, but may be attributed to clean (e.g. virus-tested) seed tuber production under certification schemes

Table 1–1 Breeding targets for the genetic improvement of potato in New Zealand

Main trait group		Traits
Agronomic	General	Marketable yield and other tuber yield components such as size, size and shape uniformity, plant vigour, and days to maturity
	Biotic	Resistance/tolerance to: late blight, potato cyst nematodes (PCN), powdery scab, common scab, aphids and viruses (e.g. PVX, PVY, PLRV), psyllid-zebra chip complex
	Abiotic	Lower water and nitrogen requirements
Tuber characteristics		Tuber size and conformation
		Skin colour and texture
		Specific gravity (dry matter content)
		Low (reducing) sugars
		Reduced internal and external defects
		Cooking qualities (e.g. sloughing, after-cook blackening)
Handling and storage		Resistance to cold-induced sweetening (CIS)
		Prolonged dormancy
		Bruising resistance
		Resistance to soft rot

(Maunder 2005), as well as to improved crop management practices, e.g. the availability of irrigation and pesticides, and cultivar development.

Breeding targets have broadened in recent years to improve the quality aspects of many food crops, such as processing qualities and nutritional values (e.g. Sands et al. 2009), but shortages of food staples and calorific malnutrition are problems still common to many of the poorest regions of the world. Productivity gains (on a per unit area basis) experienced in many global regions have not been ubiquitous (FAO 2010), yet the importance of potato as a global stable, particularly in developing countries, is increasing (Hijmans 2001; FAO 2010; Birch et al. 2012). Marketable tuber yield is an important selection target and contributes to maximising an implied economic objective in most potato breeding programmes. In many staple field crops, a high proportion of the reported yield gain (on a unit area basis) over the last century has been attributed to genetic improvement and cultivar development (Kang

2002b; Evenson and Gollin 2003), as shown in grains and oilseed crops (Duvick 2005; Mackay et al. 2011). For example, Duvick (2005) estimated that the contribution of genetics to yield gains in maize in the USA since the 1930s was approximately 50%. A more recent study on wheat, barley and canola in Europe has shown that at least 88% of the improvement in yield since the early 1980s has been due to improved genetics and the development of new cultivars, with little contribution from developments in agronomy (Mackay et al. 2011). In contrast, it has been reported that most of the yield gains in potato have been non-genetic (Sneep and Hendriksen 1979; Douches et al. 1996; Walker et al. 2003). For developing countries, it is acknowledged that some productivity gains have been achieved by breeding for disease resistances, particularly in disease-prone areas such as the sub-tropics, but the sources of yield improvement are considered to be predominantly agronomic changes and access to healthier planting material (Walker et al. 2003). In New Zealand, there is a general industry consensus based on observation that the potato crop has reached a phenotypic yield plateau in more recent years. By assessing cultivars developed and released since the 1800s, Douches et al. (1996) reported that there was little evidence of an improvement in the genetic yield potential of potato cultivars grown in the USA. This was attributed, in part at least, to greater attention being paid towards improving tuber appearance, potato processing qualities, and earlier maturity.

Quality characteristics for processing and table markets

Quality characteristics are broadly defined as those traits that determine the degree to which tubers are fit-for-purpose for their intended end-use (Simmonds 1979). As well as agronomic yield, the development of processing cultivars is concerned with increasing factory yield and reducing factory costs which can be affected by tuber quality traits. Figure 1-1 presents a general illustration of French fry manufacture, and the losses commonly incurred by the factory process. For example, skin disorders such as powdery and common scab can increase unwanted peel loss and bruising can also result in factory yield loss. Tuber shape can also affect factory yield by increasing the amount of unwanted peel and number of off-cuts. Tubers with excessive sugars, possibly as a result of cold-induced sweetening during cool storage, require a costly blanching phase (as excessive reducing sugars result in poor and often unacceptable fry quality). Increasing tuber dry matter content increases factory product yield but also reduces the amount of oil required to remove water in the frying process (Lulai and Orr 1979), although there is an optimum, as higher dry matter content can affect the

texture and palatability of potato. Somsen et al. (2004) have developed a model to predict the factory yield of potato tubers processed into French fries.

In contrast, the criteria set by the fresh potato market compared with those of the processing industry can be ambivalent and more subject to change because of consumer preferences and fashions, e.g. tuber skin and flesh colour, tuber size. Changes are usually relatively slow but it is difficult for breeders to take a long-term perspective and to define a clear breeding objective for this market sector. Furthermore, the table market itself is reasonably diverse and may, for example, include cultivars suited to washing or brushing only (which is determined by the quality of skin finish), or demand more niche cultivars such as tubers with novel skin and flesh colours (e.g. from the presence of carotenoids or anthocyanin pigments).

Abiotic and biotic considerations

Anthropogenic inputs in agricultural systems, such as nitrogen and phosphorus, and other management practices have helped to drive the increases in yield but have often been shown to have a detrimental impact on both agro- and natural ecosystems, resulting in undesirable and costly outputs, such as soil erosion and the loss of nutrients to aquatic systems (e.g. Cameron et al. 2013). There is a growing expectation in many regions that yield improvements should be achieved whilst constraining inputs, such as water and nitrogen, to improve the efficiencies and accountabilities of cropping systems and to deal with the uncertain consequences of human-induced climate change (Hijmans 2003; IPCC 2007). Input and environmental costs therefore mean that water and nutrient-use efficiencies are traits that are receiving more attention in plant breeding programmes for many field crops, including potato. This, together with finite and diminishing land resources and predicted global population growth (UN Secretariat 2011), ensures that genetic 'true productive' yield improvement is likely to remain a major focus in potato breeding programmes globally.

Biotic stresses have always been an impediment to productivity and attention is focussed on breeding for disease resistances that cause both agronomic and quality production losses. In intensive production systems, there have been financial incentives to develop agrochemicals that help to control pest and disease incursions. Their use, however, is becoming increasingly scrutinised because of environmental and food safety concerns. Further, the suite of available chemicals is sometimes ineffective against certain diseases, or the technology is impractical or just too costly to access and breeding remains the most pragmatic means to tackle these problems in the long term. Potato is susceptible to a number

of soil-borne diseases, such as potato cyst nematodes (PCN) and powdery scab, which are difficult to control by chemical means and have relied on control through resistance breeding and crop management e.g. long crop rotations. Potato cultivars are grown under conditions ranging, for instance, from temperate Asia to the subtropical lowlands and to the highlands of the Latin American Andes. Selection for this broad range of environments is reflected in the breeding populations of the International Potato Centre (CIP) in Peru. The mission of CIP is to help developing countries to add both commercial and nutritional value to tuber crops including potato and sweet potato (*Ipomoea batatas*). CIP is involved in advancing breeding populations that are suitable for a number of diverse environments across various latitudes. Therefore, breeding populations have to be adapted to a target climate with its associated biotic stresses as well as adapted to the particular regional daylength. At CIP there are a number of breeding populations undergoing recurrent selection to develop varieties that are highly resistance to late blight and potato virus X and Y (PVX and PVY). With predictions for future global climate trends (IPCC 2007), there is a greater urgency to widen adaptability to deal with the dynamic threat of biotic stress as well as tolerating climate-related abiotic stress, e.g. extreme temperature fluctuations.

Enhancing the nutritional value of potato tubers

As well as the need to maintain increases in crop yield, there is a growing awareness of the potential to enhance the nutritional quality of food staples by plant breeding, known as biofortification (Hirschi 2009; Sands et al. 2009). Malnutrition continues to be a problem of global significance, particularly in developing countries, and micronutrient deficiencies such as vitamin A, iron (Fe) and zinc (Zn) are debilitating for health and general well-being for an estimated three billion people worldwide (Khush 2008). Biofortification of staple food crops is seen as a cost-effective approach to tackle this problem without compromising agronomic productivity (Nestel et al. 2006), and has become a primary breeding target for CIP. The potato is regarded as a food crop that has good nutritional value and because of its importance as a global staple in Asia, Sub-Saharan Africa and Latin America, there are obvious benefits of further enhancing this value. Although agronomic and post-harvest practices can affect nutritional content, there is much potential to improve the health benefits of potato further by plant breeding for enhanced nutrient and phytochemical content, particularly some minerals, such as zinc and iron, and antioxidants, such as phenolics and carotenoids (Bonierbale et al. 2007; Brown 2008). Both conventional breeding and biotechnology are approaches that can be taken to improve the nutritional value of potato.

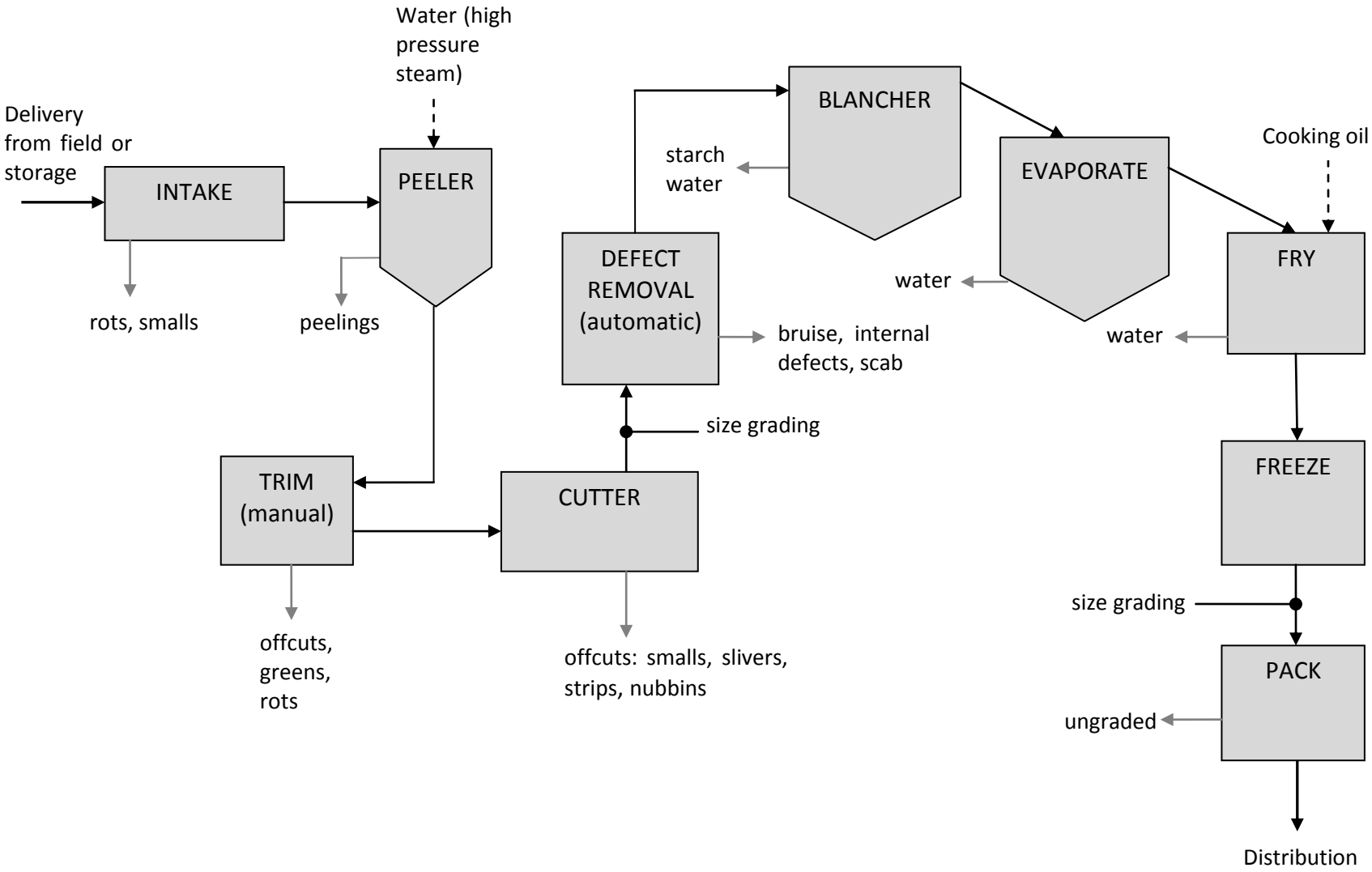

Figure 1–1 Simplified diagram of potato processing for French fry manufacture; product flow (black arrows), inputs (dashed) and losses (grey)

1.4 Breeding strategies for cultivar development

The challenge of developing new cultivars

Potato breeders are faced with meeting the high expectations of producers and consumers for cultivars to satisfy a wide and varied range of demands. In pursuit of cultivar improvement, potato breeding faces numerous technical challenges (Ross 1986). The limited contribution from breeding that has been reported for increases in tuber yield, for example, may be due to a number of possible genetic and strategic factors including (but not limited to): a narrow genetic base and limited accessibility to available genetic variation; the non-disomic mode of inheritance, complexities of non-additive gene interactions and deleterious gene load; and inefficiencies of early-stage selection, the requirements of 'maintenance' breeding (i.e. to established pests and diseases), reacting to new incursions of pests and diseases (e.g. Liefting et al. 2008), the slow multiplication rate and relatively long generation intervals. The displacement of older potato cultivars by the adoption of new genotypes with greater production potential is also slow and may, to some degree, explain the lack of progress in tuber yield production. This conservatism is not restricted only to New Zealand but is typical of potato production systems in general (Tarn et al. 1992; Walker et al. 2003; Veilleux and De Jong 2007). Developing suitable cultivars for wide-scale deployment that meet the demands of growers and consumers is both slow and resource-intensive, occurring within the constraints of a typical multi-trait, multiple-stage selection programme.

Selection methods

When the breeding objective is defined, a potato breeding scheme begins with the evaluation and selection of parental material, the crossing of the selected parents and the selection of elite clones from these progeny of crosses for further testing and potential release as cultivars. A cycle is complete when elite lines are introduced as parents for the next cycle. Population improvement by recurrent selection is therefore combined with varietal development of elite clones (Fig. 1-2). If incompatibility or sterility barriers can be overcome, germplasm resources using related species can be used as donor relatives to provide useful traits for introgression using various manipulative techniques into recipient species in a conventional pre-breeding strategy. This new germplasm may possibly undergo

several cycles of recurrent selection and then be introduced as parents in the main breeding population undergoing multiple-trait recurrent selection and deployment of advanced lines.

The New Zealand potato breeding effort is typical of many programmes worldwide whereby cultivar development traditionally uses a phenotypic-based selection strategy. Parents are chosen on the basis of their own performance or from intuition and experience of their worth from previous successes as parents. This knowledge is gained gradually by the breeder as progeny flow through the programme (Fig. 1-3) but traditionally, under this scheme, the value of a parent is not formally measured, e.g. from progeny testing, to obtain general combining ability (GCA) or breeding value estimates. Multiple crosses are made between selected parents and seedlings are grown as individual spaced plants in the field or individual pots in the glasshouse, which is common practice in many programmes for the seedling generation. This is followed by one or two stages of visual mass selection of clonal, unreplicated plots. At these initial stages, individual plants and clonal plots are selected based on their appearance or 'general worth' by the breeder. The selections are then carried forward through several clonal stages of replicated trials, with further selections made at each stage from measurement and formal statistical analysis for numerous traits.

There is a general consensus from a number of empirical studies, e.g. Anderson and Howard (1981), Tai and Young (1984), Brown et al. (1984, 1987b), Caligari et al. (1986), Gopal et al. (1992, 1994), that visual mass selection of individual plants at the early stages in the breeding cycle is ineffective, correlating poorly with subsequent clonal performance for a number of traits, including yield. From this work, an alternative strategy was proposed and subsequently developed as a recurrent genotypic selection scheme (Caligari 1992; Bradshaw et al. 2003). Progeny testing is used to identify the best families and the most promising parental material for use in crossing i.e. those parents that have high general combining ability or breeding values. Poorly performing full sib families are discarded early on in each cycle as there is a low expectation of obtaining high ranking individuals from within these families (Simmonds 1996), as developed empirically by Brown et al. (1987a) and Bradshaw et al. (1998). The best clones from within the best families are identified from further replicated field trials. These selections can be taken forward as potential cultivars for commercial deployment and/or used as parents in the next cycle. This reduces cycle time and is expected to increase the rate of genetic gain in the breeding population. Selection of superior parents can be based on their GCA from progeny tests, or mid-parent values or other cross prediction methods for untested clones (e.g. Caligari and Brown 1986; Brown et al. 1988). In the population improvement programmes at CIP, Mendoza (1989)

reported a significant response to selection for yield and some pest and disease resistances from using progeny testing compared with phenotypic recurrent selection. Progeny testing continues to be an important strategy in the CIP programme to permit (i) the selection of parents with high GCA for use in various breeding populations, and (ii) the increased chance of identifying cultivars in advanced selection cycles when true seed progenies (TPS) are provided to national programmes of developing countries (M. Bonierbale, CIP, personal communication).

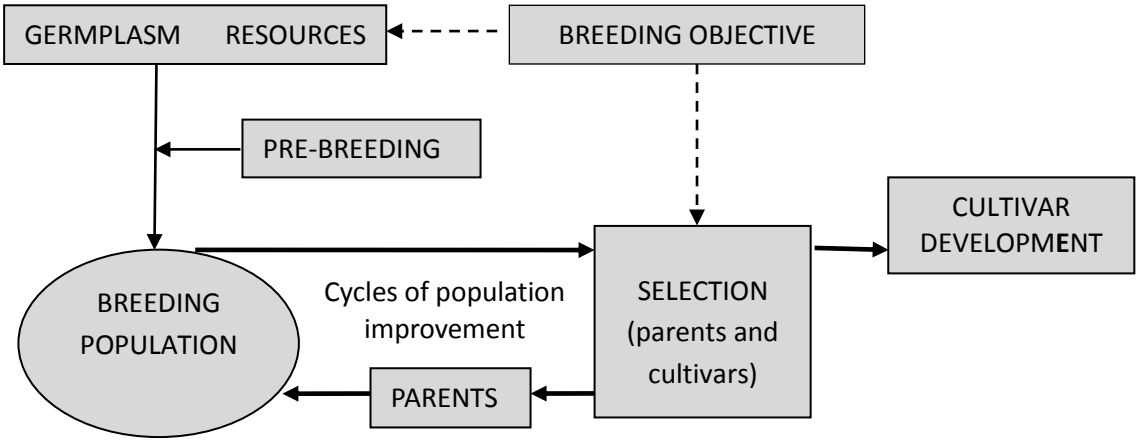

Figure 1–2 Generic illustration of a potato population improvement and cultivar development programme

Despite the reported benefits of progeny testing in breeding programmes (Mendoza 1989; Bradshaw et al. 2009), there is little evidence from literature and web-based searches that 'genotypic' recurrent selection (using progeny information to select parents) has been widely adopted as a routine system in potato breeding. This might be due to the cost (perceived or otherwise) of implementation of such a procedure or the reluctance of breeders to shift from traditional methods to a strategy which is less simple to operate. Gopal (2006), accepting that many programmes are still based on the early phenotypic discrimination of individual plants, made several recommendations for improving their efficiency, based on empirical evidence from various studies. These included: (i) the rejection of seedlings with low vigour prior to transplanting in the field; (ii) rejection of undesirable clones with characteristics that have shown a high repeatability over generations (e.g. tuber colour, tuber shape, eye depth and cracking); and (iii) negative selection initiated for tuber yield and tuber weight from the first clonal generation and number of tubers from the second clonal generation. Be it phenotypic or genotypic recurrent

selection, cultivar development is still a slow and laborious process and plant breeders and geneticists continually strive to seek new approaches to increase the efficiency of selection.

Stage	Number (approx.)	Replication	Details
Seedling	12,000	1	Single plants field grown from 'true' seed
C1 (First clonal stage)	1800	1	Four-tuber plots (unreplicated 'visual' plots)
C2	600	2	Replicated trial at Pukekohe Potato cyst nematode (PCN) testing Cold-induced sweetening (CIS) testing
C3	150	2	Replicated (early and main maturity) trials at Pukekohe Replicated (main maturity) trials at Lincoln Cooking tests
C4	50	3	Multi-site trials at Pukekohe, Lincoln and Manawatu
C5	15 – 20	3	Widespread regional trials (see Fig. 1-4). Late Blight field screening Powdery Scab field screening
C6	5 – 10	3	Widespread regional trials Heat treatment of advanced clones
C7	3 – 5	3	Widespread regional trials Commercial evaluation
C8	1 – 3	3	Widespread regional trials Commercial evaluation
C9	0 – 2	3	Agent tender for commercial release Plant Variety Rights (PVR) application

Figure 1–3 Stages of the New Zealand potato selection scheme at Pukekohe.
Note that the selection scheme at the Lincoln site in the early stages differs slightly, so that the seedling stage is grown in a glasshouse and the C1 stage is made up of field-grown single plant plots

Molecular strategies

Recent advances in biotechnology, gene discovery and molecular marker development have also aimed to accelerate the development of improved crop cultivars. These methods include transformation technologies such as trans-, cis- and intragenics, e.g. Conner et al. (2007), Jacobsen and Schouten (2009), Visarada et al. (2009); marker-assisted selection (MAS) from linkage and genome-wide association studies (GWAS), e.g. Eathington et al. (2007), Moloney et al. (2010), Schultz et al. (2012); and genomic selection (GS), e.g. Heffner et al.(2009), Jannink at al. (2010), Jonas and de Koning (2013). GS is a form of MAS that exploits dense genome-wide coverage of molecular markers to estimate all marker effects across the genome simultaneously (Meuwissen et al. 2001). It is being

applied in animal breeding programmes in New Zealand and overseas (Hayes et al. 2009a) and has caught the attention of plant breeders given the number of studies that are currently being published on the subject in relation to crop improvement. These techniques offer a means to complement and potentially to become a routine part of conventional 'field-based' crop genetic improvement methods. To date, the routine application of MAS is reported for a limited number of (mostly disease) traits (e.g. Schultz et al. 2012), but a better understanding of the genetic architecture of complex traits post-sequencing, alongside the continual reductions in costs for genotyping and automated high-throughput techniques, are likely to make MAS less prohibitive for plant breeding programmes in future (e.g. Slater et al. 2012).

A collaborative project involving thirteen countries has resulted in the publication of the potato genome sequence (Xu et al. 2011). As a result, it is important for breeding programmes to be in a position to capture the value of this work as efficiently and effectively as possible. Potato improvement programmes will have access to an increasing amount of information from various sources in a post-genomic future, which will include molecular marker and gene sequence data to complement phenotypic and pedigree information. To determine the best approach to incorporate these data into plant breeding programmes, Hospital (2007) has suggested that each crop should be treated on a case-by case basis, as the optimal strategy will depend on the peculiarities of the species, breeding objectives and traditions. Milbourne et al. (2007) concluded that '*the full potential of MAS in potato breeding may require a redesign of breeding programmes and an adoption of a marker-led ethos rather than adaptation of marker technology to present phenotype-based breeding strategies*'. As molecular understanding develops, selections will be made with an increasing abundance of both phenotypic and molecular data to hand. To exploit this information, the development of more sophisticated informatics and statistical tools for decision support will be required by the breeder.

1.5 Potato genetic evaluation and selection

Although modern genomics offers great promise for accelerating genetic gain and the rate of cultivar development, the investigation of field-based approaches to improve the efficiency of conventional selection methods should not be neglected. In more recent years, attention has been concentrated on advancing molecular-based selection methods; arguably, there has been inertia in the research and development of conventional, field-based breeding

strategies in potato. Potato breeders should continually strive to improve the efficiency of their selection methods and seek to improve the effectiveness of their breeding strategies. The development of methods to improve the effectiveness of selection has been a constant feature in crop breeding research over many years, such as field trial design (e.g. Basford et al. 1996; Edmondson 2004), the estimation of genetic parameters (e.g. Kearsey and Pooni 1996; Bernardo 2002), and the genetic evaluation of field trial data, including data collected over multiple locations (e.g. Kang and Gauch 1996; Smith et al. 2005; Crossa 2012). The motivation for breeding studies to improve selection efficiency and programme design has revolved around the much cited breeder's equation to predict genetic response: $R = \frac{i\sigma_H r_{IH}}{L}$, where i is the selection intensity, σ_H is the standard deviation of the breeding objective, r_{IH} is the selection accuracy (the correlation between the selection index and breeding objective) and L is the generation interval. There are different forms of the breeder's equation (e.g. those that acknowledge the different contributions from male and female pathways), but by expressing the selection response in this simple way, the importance of these different components and the interactions between them can be easily put into perspective. For example, the earlier selection of clones as parents will lead to a reduction in generation interval (L) but also implies lower selection accuracy (r_{IH}) because less information is available on selected parents. The investigation of methods and approaches for the evaluation of potato in a breeding programme may provide opportunities to improve selection efficiency and therefore the genetic response to selection, and is the central theme of this thesis.

Estimation (prediction) of genetic and breeding values

In recent years, the linear mixed model, with its resulting best linear unbiased predictors (BLUPs), has become the accepted method for the genetic evaluation of livestock and trees (Mrode 2005; White et al. 2007). The linear (univariate or multivariate) mixed effects model is given by:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where $\mathbf{y}$ is a $n \times 1$ vector of phenotypic observations, β is a $p \times 1$ vector of fixed effects, $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$ and $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$ are uncorrelated (respectively $q \times 1$ and $n \times 1$) random vectors, $\mathbf{X}$ and $\mathbf{Z}$ are known ($n \times p$ and $n \times q$) incidence matrices, and $\mathbf{G}$ (variance due to random genetic effects) and $\mathbf{R}$ (variance due to residual effects) are ($q \times q$ and $n \times n$) variance-covariance matrices. The elements of the vector of random effects ($\mathbf{u}$) are usually

additive genetic effects. The mixed model equations (MME) of C.R. Henderson (citations in Lynch and Walsh 1998) are used to find the solutions of β and $\mathbf{u}$:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z + G^{-1}} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \end{bmatrix}$$

Rather than pre-adjusting the phenotypes, the technique simultaneously estimates the fixed effects $\hat{\beta}$ by the generalised least-squares estimator of β (best linear unbiased estimation or BLUE) with $\hat{\mathbf{u}}$ best linear unbiased predictor of $\mathbf{u}$, e.g. the BLUPs of the breeding values. (Lynch and Walsh 1998). Typically, it is usual to speak of the 'estimation' of breeding values but strictly speaking, they are 'predictions'. BLUPs are shrunk, i.e. regressed, towards the population mean (Robinson 1991; Smith et al. 2005; Piepho et al. 2008a), so that increasing the amount of information (e.g. from relatives, repeated measures or replicated plots) will result in genotype predictions that are closer to their true values, i.e. more accurate. For random genetic effects, pedigree information can be exploited via an additive (numerator) relationship matrix **A** (Henderson 1976), with elements comprising twice the coefficients of coancestry (which are the probabilities of individuals carrying copies of the same allele by descent), which is multiplied by the genetic variance, so that:

$$\text{var}(\mathbf{u}) = \mathbf{G} = \mathbf{A}\boldsymbol{\sigma}^2_{\mathbf{u}}$$

The result is that a breeding value is fitted for all members of the pedigree, even those without trait records, hence such models being named 'individual models' (also commonly referred to as 'animal' or 'individual animal models'), as opposed to the more traditional family-based models. Individual models do not demand formal mating designs, but can deal with trials that have a variety of family structures, e.g. full-sibs, half-sibs, clones, etc., all of them accounted for via the pedigree. Non-additive genetic effects can also be included e.g. epistasis or dominance effects (Mrode 2005), which has been reported for the analysis of wheat (Oakey et al. 2006; Burgueño et al. 2007), sugarcane (Oakey et al. 2007) and canola (Beeck et al. 2010). For clonal crops, such as potato, estimates of both additive and non-additive variance components may be desirable if both breeding values for parent selection in population advancement and the total genetic value of clonal lines for cultivar selection and deployment (to exploit non-additive effects of clones) are to be evaluated.

The success of BLUP in genetic improvement schemes has been attributed to its effectiveness in disentangling management/environment effects and genetic (random) effects, and is receiving an increasing amount of attention in multi-environment trial (MET) evaluation of plant breeding data for the selection of parents and cultivars (Smith et al. 2005; Piepho et al. 2008a). Combining sources of information, such as repeated records and

relationship information, greatly increases the accuracy of selection, especially for traits with low heritabilities (Lynch and Walsh 1998). In soybean, Panter and Allen (1995a, b) found that cross prediction using BLUPs (the average of the two parents) was more closely correlated with actual performance than cross prediction using mid-parent values (mean of the two parental least-squares means). Individual models have been investigated for a diverse range of crops, including inbreeding crops such as wheat (Oakey et al. 2006), hybrids such as maize (Bernardo 1996a, b), outcrossing clonal species such as sugarcane (Chang and Milligan 1992; Balzarini 2001; Oakey et al. 2007) and potato (Tai et al. 2009), and perennials such as hops (Beatson and Alspach 2009), apple (Kumar et al. 2010) and mango (Hardner et al. 2012). Despite this, the routine adoption of breeding value estimation using linear mixed models is reported to have been slow in many crop breeding programmes (Piepho et al. 2008a). From literature searches, current routine evaluation methods in potato breeding programmes do not appear to use all available location, year and relationship information effectively and this therefore suggests an inefficient use of data. A preliminary study (Kerr et al. 2009) described the development of commercial evaluation software adapted from software originally developed for sheep and tree breeding programmes. The further development of models for evaluating potato data is required by investigating trial heterogeneity that accounts for localised spatial effects, and genetic heterogeneity by testing different (co)variance structures when analysis is extended to multiple years and locations (multi-environment trial or MET data).

Variance component estimation

The prediction of breeding values from solving the MME requires knowledge of variance and covariance components and therefore the estimation of variance components and breeding value prediction are inextricably linked. Knowledge of the genetic parameters of traits, such as heritabilities and genetic correlations, are also required to help guide an effective breeding strategy. In practice, the true variance components are unknown but are estimated from the data. Because animal and plant breeding data are generally unbalanced and models usually contain a number of nuisance parameters, maximum likelihood (ML) methods are preferred over ANOVA-based approaches. The residual (or restricted) maximum likelihood (REML) procedure (Patterson and Thompson 1971) is widely used routinely to estimate variance components and genetic parameters in plant and animal breeding data. It involves a tandem iterative procedure, so that the (genetic and phenotypic) variance components are estimated and the MME are solved for the fixed effects ($\hat{\beta}$) and

predicted breeding values ($\hat{\mathbf{u}}$); and estimates of the variance components are updated by the predicted breeding values. The cycle continues until there is convergence of the genetic and phenotypic parameter estimates. Both ANOVA and ML-based methods assume that individuals are a random sample from the population, which is unrealistic for breeding populations under selection. ML approaches can account for selection provided information that has contributed to selection decisions is included in the analysis (Piepho and Möhring 2006).

Alternatively, Bayesian inference using Markov chain Monte Carlo sampling methods (often via the Gibbs sampler) can be applied to both Gaussian and non-Gaussian traits for genetic parameter estimation and genetic evaluation (Sorensen and Gianola 2002). It is appealing for variance component and variance ratio estimation, as the posterior distribution provides confidence limits (the credible interval) as a measure of uncertainty around the point estimate (Waldmann and Ericsson 2006). The distribution of REML (co)variance estimates is unknown and only approximate confidence intervals can be calculated (Dieters et al. 1995) or, alternatively, estimated from parametric bootstrapping. Prior information can also be included in Bayesian inference if available from previous studies e.g. evaluation of previous breeding generations. A limited number of plant breeding programmes have reported the use of these approaches for genetic evaluation, possibly because of a lack of user-friendly software to apply individual models to crop data. Software such as WinBUGS (Lunn et al. 2000) and MTGSAM (Van Tassell and Van Vleck 1996) have been used for tree and crop breeding data to estimate quantitative genetic parameters, e.g. Waldmann et al. (2008) in Scots pine, Gonçalves-Vidigal et al. (2008) in common bean; and MCMCglmm-R (Hadfield 2010) for wood quality in *Pinus radiata* (Apiolaza et al. 2011).

Spatial models

Potato production is reported to be particularly sensitive to environmental variables such as the chemical and physical properties of soil (e.g. Redulla et al. 2002; Po et al. 2010). Standard blocking procedures, used in field trials, such as classical randomised complete blocks or more advanced incomplete block designs, such as row-column arrangements (Basford et al. 1996), attempt to account for such trial heterogeneity. Grouping plots into blocks assumes greater homogeneity within blocks than across the entire trial. A number of statistical approaches have been developed that deal with the presence of spatial dependence in field trials on a more localised scale (Gleeson 1997; Edmondson 2004). These are based on observations that residuals of neighbouring plots are often more alike than those of non-

neighbours. Therefore, to augment blocking and randomisation in overcoming within-trial heterogeneity or even to replace block designs altogether, a number of 'spatial' or 'neighbour' methods have been explored in order to take advantage of any spatial correlation and to exert a finer control to increase the precision of treatment estimates (e.g. Wilkinson et al. 1983; Besag and Kempton 1986; Williams 1986; Cullis and Gleeson 1991; Gilmour et al. 1997; Gleeson 1997; Piepho and Williams 2010). Studies have shown that spatial models can greatly increase the precision of genotype estimates, particularly in cereals (Gilmour et al. 1997; Qiao et al. 2000). Such improvements are not ubiquitous, however, as Müller et al. (2010) found that a standard block model outperformed a spatial model in most cases when analyzing sugar beet and barley trials. Sarker et al. (2001) recommended that block design methods could be enhanced but not replaced with spatial methods. For the fitting of spatial models, mixed models can be extended to account for correlated errors as spatial effects in an effort to account for local trends and improve estimates of genotype effects (Gilmour et al. 1997) and may result in less biased estimates of genotype effects (Smith et al. 2001). There is limited information on spatial effects in potato breeding trials.

Evaluation of multi-environment (MET) trials

Multi-environment trials (MET) generally comprise a series of trials over multiple years and multiple locations, exposing genotypes to both temporal and spatial variation in biotic and abiotic conditions, to identify superior individuals. When testing selection candidates over multiple environments, the main genotype effect and the effect of a genotype-by-environment interaction can confound the accurate discrimination between genotypes (Bos and Caligari 2008). The test locations should therefore be determined by the target end-use of improved cultivars, and may be broad or defined as being more specific to, for example, a particular region, a set of climatic variables, or certain biotic characteristics. The reality of both seasonal and location variation is such that the actual extent of MET testing will only represent a small set of the possible G×E space (Messina et al. 2009); the testing regime is ultimately determined by the resources available, with a compromise made between the costs of MET testing, the reliability of genotype value predictions and the time lag associated with achieving a desirable degree of selection precision. A better understanding of G×E effects within a MET testing regime allows a more appropriate choice of model for genetic evaluation, and a review of resource allocation and selection strategy in a breeding programme.

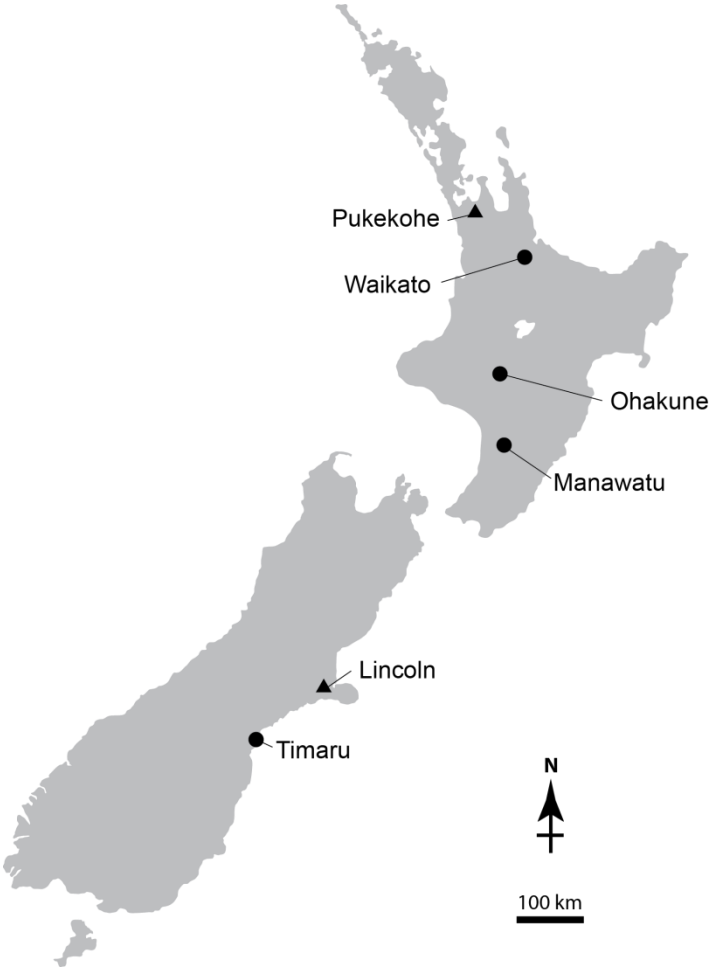

Figure 1–4 Main (▲) and regional (●) locations for potato variety trials in The New Zealand Plant & Food Research breeding programme. Pukekohe, Manawatu and Lincoln are research sites and Waikato, Ohakune and Timaru are on-farm trial sites

There are numerous statistical approaches to model G×E effects in plant breeding and, in general, these are based on univariate or multivariate methods that vary in their degree of complexity and the information that they provide (e.g. Fox et al. 1997, p.137). Over recent years, the popularity of more flexible multivariate multiplicative methods has increased, such as the 'additive main effects and multiplicative interaction' (AMMI) model (Gauch and Zobel 1988; Crossa et al. 1991). Singular value decomposition is carried out on the matrix of the two-way table of G×E effects, whereby each is modelled as the product of a genotypic score and an environmental score (or loading), and extra multiplicative (bilinear) terms are added if they improve model fit. AMMI is classified as one of several types of general linear-bilinear model (e.g. Crossa and Cornelius 2002; Yang et al. 2009), and commonly used for biplot analysis. This uses the rotated loadings of the principal components to help to simplify and interpret genotype performance and environmental

relationships, or so-called 'which-won-where' patterns, in MET analysis. Biplots, or variations thereof, provide diagnostics that can be powerful aids for understanding the relationships between test locations for planning selection strategies and for recommending cultivars, as useful information can be extracted from complex multidimensional G×E data (Fox et al. 1997), such as the delineation of groups of environments and the identification of specifically adapted cultivars. Yang et al. (2009), however, warned against their over-interpretation and misuse, particularly when uncertainty measures, such as confidence intervals, are not provided.

As an alternative to AMMI, the multiplicative mixed modelling (MMM) approach using factor analysis has been used to evaluate MET data. It is considered another class of linear-bilinear model and a mixed model analogy of the AMMI fixed-effect model (Piepho 1997, 1998; Smith et al. 2001; Smith et al. 2005). Heavy attrition of breeding lines at each stage of a MET series of breeding trials is typical of plant breeding programmes and the incomplete nature of such data is better dealt with by REML-based procedures. Furthermore, there has been a growing trend amongst crop breeders, following their animal and tree breeding counterparts, to treat genotypes as random effects, at least in the early stages of trials. The shrinkage of genotype value predictions towards the mean to allow for the uncertainty surrounding the distribution of random effects and the flexibility in analyses with, for example, inclusion of a relationship matrix or spatial components (Crossa et al. 2006; Oakey et al. 2006; Piepho et al. 2008a) makes this approach particularly attractive for the comprehensive analysis of MET data, particularly in the early stages of selection.

For the analysis of potato trial data, it is typically implied that there is no genetic covariance between sites and years, i.e. each trial is analysed independently, and genotype performance is averaged across trials. This represents an inefficient use of data. Linear mixed models can provide an improved representation of the underlying random and error components (Oakey et al. 2007). This is not only the ability to model different (co)variance structures when relationship information is used, but also to incorporate multiple traits and multi-environment (MET) trials. When genotypes are considered random, a multiplicative mixed model analysis can accommodate genotype by trial heterogeneity of variance, the correlations among genotype by trial interactions, and appropriate error variance structures for individual trials (Kelly et al. 2007). An unstructured form of the genetic variance-covariance matrix in the analysis of MET data attempts to capture the underlying genetic structure fully but is often impractical because of computational issues, particularly if there are a large number of environments or a lack of data to estimate the parameters reliably

(Kelly et al. 2007). Potato breeding data often comprise a number of relatively small trials over different years and locations where test entries appear only once. To overcome these difficulties, the factor analytic (FA) model is an approximation to the unstructured genetic variance-covariance and has been proposed as a more parsimonious form (Piepho 1998; Smith et al. 2001). FA models have been shown to model the main effects of genotypes and G×E efficiently, giving the lowest standard error of the BLUPs (Crossa et al. 2006), and have been applied to a number of crops for MET analysis, including sugarcane (Oakey et al. 2007), wheat (Crossa et al. 2006; Oakey et al. 2006), lupin (Stefanova and Buirchell 2010) and canola (Beeck et al. 2010). Further work is required to characterise the G×E component of traits and to determine appropriate evaluation models for the analysis of potato MET data.

Field trials for early stage testing

The consequence of the relatively low multiplication rates of potato tubers is that replicated clonal trials in the early stages of a potato breeding programme do not usually begin until the third or fourth year after initial crossing and longer still when deploying replicated test entries into multi-location trials (Fig. 1-4). Replication increases the precision (accuracy) of estimates of genotype differences and provides a measure of that precision (Kempton and Gleeson 1997), but some researchers have questioned the benefits of replicated trials under certain circumstances, i.e. when the ratio of genetic to phenotypic variance is reasonably large (Bos 1983a; Gauch and Zobel 1996; Bos and Caligari 2008). At the early stages, the objective is to rank a large number of genotypes for selection, rather than to predict the yield and other characteristics of a small number of potential cultivars, as is the case at more advanced stages, when more detailed information on the absolute differences between cultivars is required (Smith et al. 2005). Unreplicated trials are therefore of interest to plant breeders as they offer the first opportunity to test lines for quantitative characters in the early stages of a breeding programme before there are adequate numbers of seed or seed tubers available for planting in replicated trials: a means of foregoing selection precision for an opportunity to increase selection intensity and/or decrease the generation interval from a fixed number of test plots. Further, for multi-environment trials, the precision of across-trial comparisons is compromised in the presence of G×E effects. The magnitude of estimated average genotype × environment variances for yield in a number of crops, including potato, has previously been reported to be equal to and sometimes greater than the within-trial plot

variances (Talbot 1984). The efforts expended on maximising selection precision by replication on an individual trial site may therefore be wasted (Kempton 1984).

Unreplicated trials are often augmented by several check or standard control varieties (the behaviour of which is usually well characterised by the breeder). These are planted within each complete block to provide a measure of field variability. Unreplicated entries can be assessed against the local check, or the mean of neighbouring checks, or the weighted mean of checks by distance, for example. An advantage of using checks is that genotypes may not have to be randomised across the trial (Kempton and Gleeson 1997). In the first clonal (C1) generation of the New Zealand potato breeding programme (Fig. 1-3), unreplicated four-tuber plots are planted in non-randomised family groups with two check varieties, and selection of plots is based on visual appraisal. This approach is preferred by some breeders because of the ease with which visual within-family comparisons can be made. The common assumption is that checks will behave in a similar manner to the test entries e.g. that they are not more susceptible to the presence of a particular disease than the test entries. If this does not hold, then adjustment of test lines by the nearest checks may increase rather than decrease error (Kempton and Gleeson 1997). Augmented field designs, first proposed by Federer (1956), are based on established field designs such as the randomised complete block or incomplete block, and feature replication whereby one or more checks are replicated systematically and the remaining lines are allocated as non-replicated test entries. Again, the underlying assumption of formal augmented designs is that the checks will behave in a similar manner to the test entries (Kehel et al. 2010). A modification to established augmented designs is the partially replicated (or p-rep) design, in which standard check replicates are replaced by replicated test lines (Cullis et al. 2006; Smith et al. 2006), thus avoiding the loss of selection candidates to check cultivars. A proportion of the total number of test lines is comprised of replicated entries and allocated to plots within a formal design framework; the remaining vacant plots are subsequently filled with the non-replicated test entries. For potato, such trials may provide an opportunity to increase the number of selection candidates that are tested in a single site and also to extend trials to multiple locations for MET testing at an earlier stage than is currently practised. In MET, different subsets of genotypes can be used as replicates in individual trials across multiple locations. This could result in an improvement in the selection efficiency at the early stages of a potato breeding programme, increasing both the rate of genetic gain and variety deployment.

1.6 Study motivation and aims

From researching the literature and from various discussions with potato breeders, it is apparent that there has been limited investigation in recent years on the breeding strategy of potato, particularly on the approaches to evaluation and selection in the early stages of a genetic improvement programme. Potato breeding and the selection of quantitative traits tends to follow traditional habits, and decision-making is often based on breeder experience and practical considerations rather than any objective genetic insight, such as genetic correlations between traits or the breeding values of individuals (Vermeer 1990; Bradshaw and Mackay 1994). A study of field-based conventional selection methods in the genomics era may seem an anachronism to some, but it continues to deserve our attention, not least because the cost of running a breeding programme, particularly in terms of fixed costs, is under constant review. Increasing the effectiveness and efficiency of selection for crop improvement by influencing aspects of the breeder's equation is a central tenet of plant breeding research and this thesis is no exception. Several approaches to improve the selection strategy for population improvement and cultivar development are investigated; statistical models, evaluation methods and selection strategies, focusing on key agronomic, disease and nutritional traits, are explored for the genetic improvement of potato. Each chapter of this thesis is intended to be read as a discrete piece of research, but there are many connections between them, as each is concerned with improving the efficiencies of potato selection, particular at the early stages of the selection cycle. As a consequence, there may be some repetition of material. The structure is as follows:

Chapter Two determines the variance components and genetic parameters for important micronutrients and are investigated using univariate and multivariate Bayesian models. Micronutrient malnutrition is a global health problem particularly in developing countries. A better understanding of the genetic control of important micronutrients, such as iron and zinc, will help in the devising of breeding strategies for the biofortification of potato tubers.

Chapter Three examines the consequences of reducing the degree of replication (by partially replicating a portion of the trial) at the early stages of selection for tuber yield and yield components, using empirical data and simulation. This work questions whether the response to selection can be increased by increasing the number of test entries (and hence foregoing some selection accuracy), and distributing entries to multiple test locations at an earlier stage than is currently practised.

Chapter Four explores different variance models for the genetic evaluation of powdery scab of tubers, using long-term trial data. Various models are tested to determine the most appropriate evaluation model for the prediction of breeding values for resistance to this soil-borne disease. The data also provide an opportunity to examine the genetic trend of resistance in the breeding population and to determine the effectiveness of selection for powdery scab resistance over the term of the trials.

Chapter Five To complement the work of Chapter Four, different variance models are tested for the MET evaluation of potato yield data. Spatial models are also tested to establish if they enhance standard incomplete block designs by additional error control in the analysis of tuber yield.

Chapter Six is an extension of Chapter Five, whereby analysis using a multiplicative mixed model is extended to later-stage trials when selection candidates are tested over multiple locations in New Zealand to evaluate clonal performance and stability. The aim is to assess environments as well as genotypes in terms of G×E. The genetic improvement of commercial cultivars for tuber yield over the past 50 years is also appraised.

Chapter Seven brings this programme of research to a close by considering the implications of conclusions drawn from the work, discussing several specific issues and suggesting areas for further development.

2 Genetic parameter estimation of micronutrient traits in diploid potato from a base population of Andean landrace cultivars and the implications for breeding

2.1 Summary

Micronutrient malnutrition is a global health problem. An improved understanding of the genetic variation of important micronutrient traits within a potato breeding population will help devise breeding strategies for the biofortification of this important food staple. The dataset consisted of 556 individuals from 17 full-sib diploid families grown in 2006 in Huanuco, Peru and 1329 individuals from 32 full sib families grown in 2009 in Ayacucho, Peru. Genetic parameters were estimated using univariate and multivariate 'individual' Bayesian models for micronutrient tuber content including iron and zinc. Genetic variance was additive and heritability estimates were moderate (0.36 to 0.57) and inflated if the common environment of full-sibs was not taken into account. Posterior modes of genetic correlation estimates between minerals, when analysed on a dry-weight basis, were all positive (0.04 to 0.72) and between minerals and tuber dry matter were negative. (-0.14 to -0.38). On a fresh-weight basis, genetic correlations between minerals and tuber dry matter were small but positive (0.05 to 0.18). The implications and challenges for selective breeding to enhance micronutrient content in potato tubers are discussed.

2.2 Introduction

Improving the health benefits of major food staples by enhancing micronutrient content of essential vitamins and minerals in the edible portions has become an important target for plant breeders in recent years (e.g. Graham et al. 1999; Gregorio 2002; Nestel et al. 2006; Pfeiffer and McClafferty 2007a; Sands et al. 2009; Bouis and Welch 2010). Biofortification is the genetic improvement of the nutritional value of food crops through conventional plant breeding or biotechnology. It is supported by predictive cost-benefit analysis as an effective approach to help reduce micronutrient deficiencies (Nestel et al. 2006) and has been endorsed as a priority development goal by the Copenhagen Consensus, an international think-tank on global poverty (Horton et al. 2009). Global micronutrient deficiencies do not tend to receive the same attention (from the media or otherwise) as calorific malnutrition and are a problem in poorer communities in particular, especially for women, infants and

children. Affected communities may often have an adequate supply of carbohydrate and protein but lack some vitamins and minerals that are essential for healthy body function. The effects of micronutrient deficiencies are not always immediately apparent and therefore are often described as a 'hidden hunger' (Stein et al. 2005). Iron deficiency alone is estimated to affect 2.7 billion people globally (Hirschi 2009), and the effects are reported to include impaired physical activity, impaired cognitive development, and both maternal and infant mortality. Zinc deficiency is also a widespread global problem and can lead to infant and child respiratory infection, diarrhoea, stunting and mortality (Stein 2010).

Reducing micronutrient malnutrition is likely to lead to an improvement in public health and in economic outcomes at a local scale and beyond, as well as an improved quality of life for individuals (Stein et al. 2005). The importance of potato as a food staple in poorer regions of Asia, sub-Saharan Africa and Latin America combined with evidence for genetic variability for mineral concentrations in a favourable food matrix have made biofortification a new potential breeding target at the International Potato Centre (CIP) (Bonierbale et al. 2007). Although agronomic and post-harvest practices can affect nutritional content (Rengel et al. 1999; Hirschi 2009), the variation in micronutrient levels in many food crops is considered to have an exploitable genetic component (Graham et al. 1999; Gregorio 2002). Knowledge on the level and type of genetic variation present in crop gene pools is required to help determine an appropriate breeding strategy.

A crop breeding programme requires estimates of variance components, not only to obtain genetic parameters to help define a breeding strategy, but also to predict breeding values to identify superior parents and breeding lines for variety development. Linear mixed models provide an improved representation of the underlying random and error components, i.e. the ability to model different (co)variance structures when pedigree information is used and analysis is further extended to multiple traits and multi-environment (MET) trials (Oakey et al. 2007). Pedigree information is exploited via the relationship matrix **A** (Henderson 1976), accounting for the expected additive genetic relationships between all individuals in the pedigree. Exploiting these relationships, a breeding value can be fitted for all members of the pedigree, even those without trait records, hence such models are named 'individual' or 'individual plant' models (but more commonly referred to as 'animal models') as opposed to the more traditional family-based approaches. Combining information on the individual and all relatives in a selection programme greatly increases the accuracy of selection, (Lynch and Walsh 1998). Variance components are required for the estimation of the BLUPs (best linear unbiased predictors) of breeding values. In

practice, the true variance components are unknown but are estimated from the data either by likelihood approaches, usually by REML (Restricted Maximum Likelihood), (Patterson and Thompson 1971), or from Bayesian inference (e.g. Sorensen and Gianola 2002). Bayesian inference using Markov chain Monte Carlo (MCMC) sampling methods (often via the Gibbs sampler) is attractive for variance component and variance ratio estimation as the posterior distribution provides the credible interval as a realistic measure of uncertainty around the point estimate (Waldmann and Ericsson 2006). Prior information can also be included in Bayesian inference if available from previous studies, e.g. evaluation of previous breeding generations. Bayesian methods have remained out of reach for most plant breeders because of the apparent lack of user-friendly software to apply individual models to crop data. This may partly explain the limited number of crop breeding programmes reporting the use of these approaches for genetic evaluation.

Previous studies have demonstrated genetic diversity in Andean potato germplasm for micronutrient traits. Andre et al. (2007) found significant diversity in the tuber content of iron, zinc, calcium, vitamin C, carotenoids and phenolics from a sample of 74 genotypes of a CIP core collection, which was made up of 8 taxonomic groups from the *Solanum tuberosum* species. Burgos et al. (2007) identified genotype variability in iron and zinc concentrations for landrace cultivars from several taxa of *Solanum*. Derived from a base population of diploid landrace accessions, the breeding population of the present study was initiated in 2004 at CIP in Lima, Peru in coordination with the HarvestPlus 'Biofortification Challenge Program' (Pfeiffer and McClafferty 2007a). CIP aims to enhance the micronutrient content of potato tubers at the diploid level and use this material in a pre-breeding strategy prior to introduction as parental material into tetraploid breeding populations. The objective of this study was to estimate variance components and genetic parameters of important micronutrient traits from a breeding population based on landrace genotypes using data collected from tuber progeny field tests. This will assist in the recommendation of selection procedures and the development of a breeding strategy for biofortification. The study also illustrates that Bayesian procedures using the MCMC to fit the individual model are now more accessible to plant breeders for the routine estimation of variance components, genetic parameters and breeding values.

2.3 Materials and Methods

Plant material

Three diploid cultivar groups of *Solanum tuberosum*, namely *stenotomum, goniocalyx* and *phureja* made up the parental base population. For the first generation (G_1), a sample of cultivars of the three species from the study of Burgos et al. (2007) was identified as base parents (G_0) and crossed following a nested mating design (Table 2-1; Table 10-1 & Fig. 10-1, Appendix I), i.e. each of a group of males mated to a subset of females. 17 full-sib families and 4 half-sib families from 4 males and 16 females were generated; 703825 ('China Runtush') and 703421 ('Poluya') were both female and male parents. The first generation (G_1) was grown in 2006 in Huanuco, Peru, at an altitude of 3800m. Tuber families, consisting of three tubers (clones) per genotype, were grown within full-sib family groups with three replications of each family in a randomised complete-block design (RCBD). Planting distances were 0.3m between plants and 0.9m between rows. At harvest, tuber samples of 12 genotypes, if possible, were taken at random from each replicate within each family for micronutrient analysis.

All analyses were conducted on peeled tubers. Mineral content was determined by inductively coupled plasma-optical emission spectrophotometry (ICP) at Waite Analytical Services in Australia. For further details of tuber sample preparation and analytical methods for mineral determination, see Burgos et al. (2007). Micronutrients analysed included iron, zinc, calcium and vitamin C. Aluminium was used as an indicator of contamination of samples with soil or dust, as it is often found in higher levels in the soil and lower levels in grains and tubers (Pfeiffer and McClafferty 2007b). Ascorbic acid (AA; vitamin C) concentrations were evaluated by the spectrophotometric method of Egoville et al. (1988). The method is based on the ability of AA to reduce dye 2,6-dichloroindophenol. Concentrations are expressed in mg/100g, fresh weight. The dry matter content of the individual samples was determined on the basis of differences in weight before and after oven drying at 100°C and used to estimate the concentration in mg/100g, dry weight. In G_1, there were 556 observations, which included 487 for mineral content and 527 for vitamin C and dry matter content. Family sizes analysed ranged from 23 to 36 genotypes.

Parent selection for the second generation (G_2) was based on the phenotypic values of individuals from the G_1 trials for higher iron, zinc and other desirable agronomic characteristics. Over 40 potential parents were initially chosen, but natural attrition (due to male or female parent sterility, for example) resulted in a final crossing scheme made up of

8 female parents and 8 male parents intercrossed in a factorial mating design, i.e. each female member of the group was mated to each male member using 2 sets of 4 females × 4 males generating 32 full-sib families (Table 10-2, Appendix I). For G_2, seedlings from the factorial crosses were transplanted into the field in Huancayo, Peru (2007–2008), using a RCBD with 4 replicates and 30 plants per replicate. At harvest, a set of tuber families from across the complete trial were retained and planted as a RCBD in Ayacucho, Peru (2008–2009). Three plants (clones) per genotype were planted in each plot within full-sib family groups with 3 replicate groups per family. At harvest, tuber samples from clones of each three-plant plot were pooled and analysed for the micronutrient content of peeled tubers by ICP and for dry matter content, as previously described. In total there were 1329 progeny records analysed for iron, zinc, calcium and dry matter content, with family size ranging from 19 to 74 genotypes. There was no data collected for vitamin C from the G_2 Ayacucho breeding trial.

Table 2–1 Base parents (G_0) of the first generation (G_1)

Female	Group[†]	Cultivar name	Male	Group[†]	Cultivar name
702736	*Stn*	Puca Micnush	**703287**	*Stn*	Cceccorani
703280	*Gon*	*Unknown*			
703312	*Stn*	Morada Taruna			
703317	*Stn*	Chingos			
702815	*Stn*	Morar Nayra Mari	**703421**	*Stn*	Poluya
703291	*Phu*	Rosca			
703825	*Gon*	China Runtush			
704393	*Gon*	Maria Cruz			
701165	*Stn*	Calhua Rosada	**703825**	*Gon*	China Runtush
[‡]**703168**	*Gon*	Puca Pishgush			
703352	*Gon*	Cashpadana Amarilla			
[§]**703421**	*Stn*	Poluya			
703831	*Gon*	Pampuna			
703831	*Gon*	Pampuna	**704218**	*Phu*	Yema de Huevo
700313	*Stn*	Cuchipa Ismaynin			
703197	*Stn*	Yana Sucre			
704481	*Gon*	Amarilla			

[†]*Stn: Stenotomum; Gon: Goniocalyx; Phu: Phureja.* [‡]No progeny measured for mineral content. [§]No progeny measured for vitamin C content.

Data analysis

A Bayesian approach based on an individual model was used to estimate variance components, heritabilities and genetic correlations for various micronutrient traits in potato. The general form of the full univariate model was:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{Z_1a} + \mathbf{Z_2c} + \mathbf{Z_3f} + \mathbf{e}$$

where **Y** is a vector of observations on the trait under study and **X**, $\mathbf{Z_1}$, $\mathbf{Z_2}$ and $\mathbf{Z_3}$ are known incidence matrices. In a traditional generalized linear model, the vector of replicate effects **b** may be considered as fixed effects but in the Bayesian analysis were fitted with a prior of zero mean and large variance. The vector of random additive genetic effects of individual genotypes, **a**, has the distribution assumed to be multivariate normal (MVN), with the parameters $(0, \sigma_a^2, \mathbf{A})$, **c** is a vector of common environmental effects with the distribution assumed to be MVN, with the parameters $(0, \sigma_c^2, \mathbf{I}_c)$, **f** is a vector of family effects with the distribution assumed to be MVN, with the parameters $(0, \sigma_f^2, \mathbf{I}_f)$ and **e** is the vector of errors distributed MVN with parameters $(0, \sigma_e^2, \mathbf{I}_e)$, $\mathbf{I_c}$, $\mathbf{I_f}$ and $\mathbf{I_e}$ represent identity matrices of size equal to the number of common environments, families and plants respectively. The subscripted σ^2 is the variance of each of the random effects. **A**, the numerator relationship matrix, describes the additive genetic relationships among individual genotypes and was generated from the pedigree. In matrix format, the random effects from the general form of the univariate model are defined by:

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{c} \\ \mathbf{f} \\ \mathbf{e} \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_c^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_f^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{bmatrix} \right)$$

Data were log-transformed for both univariate and multivariate analyses. Weak priors were assumed for variance components in G_1 that followed an inverse χ^2 distribution with 1 degree of freedom, $\sigma^2 \sim Inv-\chi^2(1,\phi)$,where ϕ is a scale parameter which apportioned the prior variance equally between the variance components. Trait variances and covariances estimated from the multivariate analysis in G_1 were used as priors for the analysis of G_2 data.

In a factorial design (G_2), the full-sib family component of variance (the male × female interaction) is expected to estimate 1/4 of the dominance effect (Bernardo 2002). The σ_a^2 and σ_f^2 variance components have the following genetic expectations under the linear mixed model:

$$\sigma_a^2 = 4\sigma_{GCA}^2 = V_A + 1/4V_{AA} + 1/16V_{AAA} + \ldots$$

$$\sigma_f^2 = \sigma_{SCA}^2 = 1/4V_D + 1/8V_{AA} + 1/8V_{AD} + 1/16V_{DD} + \ldots$$

where V_A is the additive genetic variance, V_D is the dominance genetic variance and V_{AA}, V_{DD} and V_{AD} are the epistatic genetic variances due to interactions of additive effects, dominance effects and additive and dominance effects at two loci, σ^2_{GCA} is the variance due to the general combining ability (GCA) of the parents and σ^2_{SCA} is the variance due to the specific combining ability (SCA) of the crosses. For this study, epistatic genetic effects were assumed negligible and σ^2_f was not estimated from data of the nested design (G_1) due to the relatively small number of full-sib progenies measured. In general, heritability (narrow-sense) was obtained from:

$$h^2 = \frac{V_A}{V_P} = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_c + \sigma^2_f + \sigma^2_e}$$

where V_P is the phenotypic variance with common environment and family (G_2 only) components of variance included where appropriate, as indicated by deviance information criterion (DIC) tests (Spiegelhalter et al. 2002), for both univariate and multivariate analyses.

The common environment ratio of full sibs was:

$$c^2 = \frac{V_c}{V_P} = \frac{\sigma^2_c}{\sigma^2_a + \sigma^2_c + \sigma^2_f + \sigma^2_e}$$

where V_C is the common environmental effect of full-sibs.

Heritability in the broad-sense was obtained from:

$$H^2 = \frac{\sigma^2_a + 4\sigma^2_f}{\sigma^2_a + \sigma^2_f + \sigma^2_c + \sigma^2_e}$$

All models were fitted using Markov Chain Monte Carlo methods implemented in R (R Development Core Team 2012) using MCMCglmm (Hadfield 2010). For univariate analyses, 80,000 iterations were used, storing every 34th sample after an initial burn-in of 12,000. Posterior modes of variance components, and narrow-sense heritabilities from a univariate model in G_1 for all traits were reported. The model for the univariate analyses in G_1 included additive and common environment effects but ignored any full-sib family effect.

Univariate models were further extended to accommodate multivariate analyses, which included iron, zinc, calcium, vitamin C and dry matter in G_1 and iron, zinc, calcium and dry matter in G_2. For multivariate analyses in both G_1 and G_2, iteration number was increased to 250,000, storing every 95th sample after an initial burn-in of 60,000. Different (co)variance structures for the random effects were fitted, as outlined in Table 2-4, where DIAG fitted

different trait variances and zero covariances between each pair of traits and US (unstructured) fitted both different traits variances and covariances between each pair of traits:

$$\mathbf{DIAG} = \begin{bmatrix} \boldsymbol{\sigma}^2_{t_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}^2_{t_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\sigma}^2_{t_n} \end{bmatrix} \qquad \mathbf{US} = \begin{bmatrix} \boldsymbol{\sigma}^2_{t_1} & \boldsymbol{\sigma}_{t_1,t_2} & \cdots & \boldsymbol{\sigma}_{t_1,t_n} \\ \boldsymbol{\sigma}_{t_2,t_1} & \boldsymbol{\sigma}^2_{t_2} & \cdots & \boldsymbol{\sigma}_{t_2,t_n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\sigma}_{t_n,t_1} & \boldsymbol{\sigma}_{t_n,t_2} & \cdots & \boldsymbol{\sigma}^2_{t_n} \end{bmatrix}$$

so that $\sigma^2_{t_n}$ is the variance for trait *n* and $\sigma_{t_{n_1},t_{n_2}}$ represents the covariance between two traits, n_1 and n_2. Models were tested using the DIC; Models 1–5 and Models 1–8 were tested for G_1 and G_2 data respectively. A summary of the genetic parameters was provided by the mode and 95% credible interval of the posterior distributions.

Table 2–2 Summary of phenotypic micronutrient data (mg kg^{-1} dry weight) and dry matter content (%) in the first and second generations (G_1 and G_2 respectively)

Cycle	Trait	n	Minimum	Maximum	Mean	†SD	‡CV%
G_1	Iron	487	9.5	37.3	19.0	3.9	20.4
	Zinc	487	7.2	27.5	15.8	3.0	18.7
	Calcium	487	40.5	780.0	163.9	87.2	53.2
	Vitamin C	527	140.4	918.7	399.5	117.9	29.5
	Dry matter	487	15.5	36.4	26.4	2.9	11.0
G_2	Iron	1329	7.0	42.5	21.0	5.0	23.7
	Zinc	1329	2.8	38.9	15.4	3.4	22.4
	Calcium	1329	52.3	689.7	171.9	71.0	41.3
	Dry matter	1326	12.6	35.0	26.0	3.3	12.1

†Standard deviation. ‡Coefficient of variation as the standard deviation expressed as a percentage of the mean.

2.4 Results

Table 2-2 summarises the mineral and vitamin data for G_1 and G_2 on a dry-weight basis. Coefficients of phenotypic variation (CV%) for micronutrients were highest for calcium followed by vitamin C. Variation for iron and zinc were similar in both G_1 and G_2. Mean and CV% for tuber percentage dry matter content were also similar in both G_1 and G_2.

Table 2–3 Posterior modes for variance components and heritabilities from univariate analyses of first generation (G_1) data

	Variance components				Heritability
	σ^2_a	σ^2_c	σ^2_e	c^2	h^2[*95% credible interval*]
Iron	0.034	0.015	0.024	0.21	0.45 [0.30 0.65]
Zinc	0.028	0.013	0.019	0.22	0.42 [0.32 0.63]
Calcium	0.109	0.031	0.133	0.15	0.36 [0.18 0.70]
Vitamin C	0.043	0.035	0.033	0.32	0.38 [0.23 0.62]
Dry matter	0.014	0.011	0.008	0.35	0.41 [0.31 0.54]

Posterior modes of heritabilities for micronutrient traits from univariate analyses were moderate, as shown in Table 2-3 (with the 95% credible interval of the posterior distributions in parentheses). Estimates for all traits were inflated when the common environment of full-sibs was not taken into account (results not tabulated), such that heritabilities for iron, zinc and vitamin C were 0.67 [0.46, 0.79], 0.70 [0.52, 0.81] and 0.76 [0.55, 0.88] respectively. The posterior distributions of narrow-sense heritabilities for vitamin C, iron and zinc respectively are presented in Figures 2-1a, 2-1b and 2-1c.

For the multivariate analyses of G_2, eight different models were fitted (Table 2-4). Models 1 and 2 were equivalent to running univariate analyses for each trait as all traits are assumed independent with zero covariances and heterogeneous variances. Based on the DIC, Model 4 was the best fitting model; a multivariate model with unstructured (co)variance matrices for both individual (genotype) and residual error, and common environment effects with heterogeneous variances and zero covariances between response variables (traits). The inclusion of a full-sib family effect did not improve the model fit. Although Model 4 was the preferred model, the broad-sense heritabilities may be of interest and are therefore presented, with estimates (from Model 6) of 0.57 [0.43, 0.72], 0.55 [0.38, 0.69] and 0.59 [0.46, 0.74] for iron, zinc and calcium respectively. Model 4 was also the best fitting model in the multivariate analysis of G_1 (results not shown) although the family effect (Models 6–8) was not tested.

Posterior modes of narrow-sense heritabilities, as shown in Tables 2-5 to 2-7, were moderate. From G_1 to G_2, estimates increased for iron (marginal increase) (Fig. 2-1b), calcium and dry matter, and slightly decreased for zinc (Fig. 2-1c) but were relatively stable given that trials were over two different sites and years. Analysis of G_2 data was repeated using weaker priors of variance components, reducing the degree of belief. In comparing the two runs, the MCMC trace output appeared to be reasonably stable, with heritability

estimates [credible intervals] of 0.44 [0.25, 0.67] for iron, 0.30 [0.17, 0.58] for zinc, 0.60 [0.37, 0.76] for calcium and 0.27 [0.15, 0.41] for dry matter.

Posterior modes of the genetic correlations between iron and zinc on a dry weight basis were positive in both G_1 and G_2 (Tables 2-5 and 2-6, Fig. 2-1d). Genetic correlations between calcium and iron/zinc were close to zero and shifted to become more positive from G_1 to G_2, and correlations between the mineral traits and vitamin C in G_1 were effectively zero. Between the mineral traits and dry matter content, correlations were negative in both G_1 and G_2 (Tables 2-5 and 2-6, Fig. 2-1e and 2-1f). In comparison, genetic parameters estimated on a fresh-weight basis were similar in general, with the exception of the genetic correlations between the minerals and dry matter content which were positive (G_2 results shown in Table 2-7).

Table 2–4 (Co)variance structures and model deviance information criterion (DIC) for the multivariate analyses of G_2 data where DIAG has a zero covariance structure and US an unstructured covariance between the response variables

Model	Individual	Common environment	Family	Error	†DIC	
1	DIAG	-	–	DIAG	2008	2001
2	DIAG	DIAG	–	DIAG	1470	1472
3	US	-	–	US	781	781
4	US	DIAG	–	US	0	4
5	US	US	–	US	14	12
6	US	DIAG	DIAG	US	38	43
7	US	US	DIAG	US	32	36
8	US	US	US	US	31	32

†Difference in DIC (2 runs) from Model 4 (set to zero)

Table 2–5 Posterior modes for heritability (diagonal) and additive genetic correlations (below diagonal) for iron, zinc, calcium, vitamin C and tuber dry matter content from a multivariate analysis of G_1 data (Model 4) estimated on a dry-weight basis

Trait	Iron	Zinc	Calcium	Vitamin C	Dry matter
Iron	**0.41** [0.29 0.50]				
Zinc	0.45 [0.32 0.64]	**0.38** [0.27 0.47]			
Calcium	0.04 [-0.23 0.34]	0.12 [-0.15 0.39]	**0.42** [0.29 0.64]		
Vitamin C	-0.01 [-0.18 0.29]	0.10 [-0.15 0.30]	0.05 [-0.27 0.33]	**0.38** [0.28 0.50]	
Dry matter	-0.23 [-0.42 -0.06]	-0.24 [-0.41 -0.07]	-0.19 [-0.36 0.07]	-0.06 [-0.28 0.10]	**0.32** [0.22 0.38]

Table 2–6 Posterior modes for heritability (diagonal) and additive genetic correlations (below diagonal) for iron, zinc, calcium and tuber dry matter content from a multivariate analysis of G_2 data (Model 4) estimated on a dry-weight basis

Trait	Iron	Zinc	Calcium	Dry matter
Iron	**0.43** [0.28 0.65]			
Zinc	0.72 [0.42 0.88]	**0.36** [0.17 0.54]		
Calcium	0.35 [-0.04 0.61]	0.57 [0.18 0.76]	**0.57** [0.37 0.71]	
Dry matter	-0.34 [-0.61 0.08]	-0.38 [-0.66 0.10]	-0.14 [-0.49 0.20]	**0.42** [0.25 0.57]

Table 2–7 Posterior modes for heritability (diagonal) and additive genetic correlations (below diagonal) for iron, zinc, calcium and tuber dry matter content from a multivariate analysis of G_2 data (Model 4) estimated on a fresh-weight basis

Trait	Iron	Zinc	Calcium	Dry matter
Iron	**0.45** [0.27 0.59]			
Zinc	0.61 [0.33 0.84]	**0.26** [0.14 0.42]		
Calcium	0.07 [-0.32 0.52]	0.45 [-0.02 0.77]	**0.51** [0.31 0.80]	
Dry matter	0.18 [-0.13 0.36]	0.14 [-0.13 0.38]	0.05 [-0.23 0.27]	**0.52** [0.41 0.60]

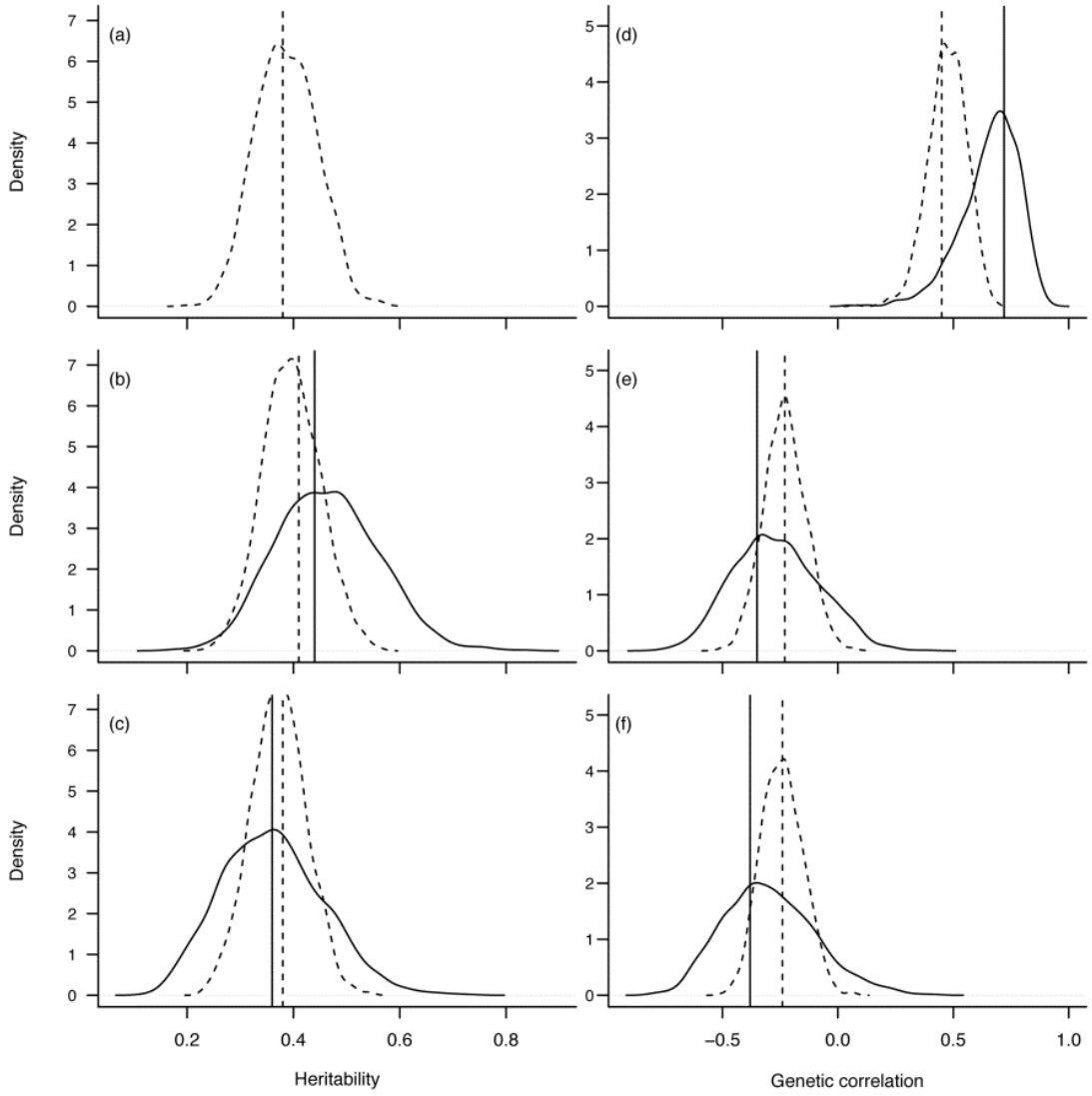

Figure 2–1 Marginal posterior distributions for narrow-sense heritability of: (a) vitamin C in G_1; (b) iron in G_1 and G_2; (c) zinc in G_1 and G_2. Marginal posterior distributions for additive genetic correlation between: (d) iron and zinc in G_1 and G_2; (e) iron and dry matter content in G_1 and G_2; (f) zinc and dry matter in G_1 and G_2; where - - - is G_1 and — is G_2 and the vertical lines indicate the posterior modes of the distributions

2.5 Discussion

Genetic variation and heritabilities

From a breeding perspective, it is acknowledged that the success of biofortification will be determined by the type of genetic control and amount of genetic variation, the relationships between the target micronutrients with other important agronomic and quality traits, and genotype stability for target micronutrients across different environments. In the present study, the additive genetic variance (V_A) was estimated directly making use of the additive

genetic relationships via the **A** matrix, and the dominance variance (V_D) from the estimate of the full-sib family effect (σ_f^2) with the expectation of $1/4V_D$. This study supports the hypothesis that tuber micronutrient content is under genetic control and in the population studied, this control appeared to be additive. There was insufficient data to reliably detect any real non-additive genetic component. Exploiting the additive gene effects present in this diploid population will result in the genetic improvement of important micronutrients in potato tubers. Furthermore, the magnitudes of narrow-sense heritability estimates point towards individual rather than family-based selection as a selection strategy for important micronutrients. The moderate heritabilities also suggest that the level of within-family variation is such that superior individuals will potentially be identified from within a number of families. Graham et al. (1999) suggested that the mechanisms controlling the uptake, transport and loading of micronutrients are likely to be additive, indicating that emphasis should be placed on an approach of population improvement from recurrent selection. Parents of the G_2 progeny were selected from individuals of G_1 that had higher iron and/or zinc but selections also included genotypes with desirable agronomic features. Truncation selection is expected to reduce additive genetic variance due to Bulmer's gametic-phase disequilibrium (Falconer and Mackay 1996). No inference can be made on this effect with these data, of course, because of the large sampling errors involved but it should be noted that selection in this case was not strictly truncated and the selection differential would have been reduced because many preferred crosses did not result in progeny. There is limited information available in the published literature about the genetic control of micronutrient content in staple crops, including potato, using data from designed crossing trials. In studies on the variation in potato clones (commercial tetraploid breeding lines and cultivars), Brown et al. (2010, 2011) estimated broad-sense heritabilities and reported significant genetic variation and genotype-by-environment (G × E) interaction for tuber iron content, but a scarcity of exploitable variation for zinc tuber content in two out of three trial locations. A study by Haynes et al. (2012b) on *S. tuberosum* x (*S. phureja-S. stenotomum*) 4x-2x clones found significant levels of genetic variation for a number of micronutrients and a G × E interaction for zinc. In an assessment of 23 potato genotypes for zinc content after applications of foliar zinc fertilizers over 4 years, White et al. (2012) identified significant genotype differences and environmental effects, but no evidence of G × E effects. Burgos et al. (2007) found significant environmental effects and G × E effects over two highland locations for iron and zinc in native Andean diploid accessions, a number of which formed the basis of the breeding

population in the present study. The G × E interaction in this case was due largely to a re-scaling of genotypes suggesting heterogeneity of genetic variance, which is of less concern than if there had been a significant re-ranking of genotypes.

Although not apparent in the present study, significant changes in heritability estimates for tuber mineral content across trials and/or years may be expected, given that various environmental components affecting crop mineral availability have been reported such as the physical and chemical properties of soil, e.g. White and Zasoski (1999); Po et al. (2010). Low heritabilities due to large within-trial error variances will compromise the selection for micronutrient traits in the early stages of a selection programme. Accumulation of zinc (and, to a lesser extent, iron) in edible portions of food crops is reported to be particularly sensitive to environmental variables (Pfeiffer and McClafferty 2007b), which suggests that more effective strategies to identify genotypes with high stable mineral expression across environments may be required. Soil zinc deficiency is common in many crop growing regions (White and Zasoski 1999; Cakmak 2008) and its availability and accumulation in the edible portions of crops is therefore likely to be a complex function of soil mineral status and interactions with other environment, agronomic and management factors (e.g. White and Zasoski 1999; White et al. 2009; Po et al. 2010). In durum wheat, for example, there is evidence that nitrogen status affects iron and zinc accumulation in grain. (Kutman et al. 2011) and nitrogen availability itself will be dependent upon soil N status and condition, amongst other interacting factors. In potato tubers, zinc assimilation in tubers has been linked with the co-transport of nitrogen (White et al. 2012), but in long-term field studies, Šrek et al (2010) found no differences in the iron and zinc content of tubers under different rates of nitrogen, potassium and phosphorus field treatments. Further research on the extent and type of genotype-by-environment interactions of key micronutrients in potato would give some indication of the requirements for MET testing and the benefits of developing a marker-assisted selection strategy.

In the RCBD, replication of full-sib families should go some way at least to remove confounding that occurs between the common environment effect and dominance effects. Failing to account for non-additive and non-genetic effects can inflate heritabilities and reduce the precision of breeding value estimates. The grouping of families is common in early-stage trials and is often preferred by the breeder for practical reasons e.g. visual evaluation of groups of individuals with known parental combinations. A common environment effect was fitted in the analysis which appeared to account for some bias in the additive variance estimates; ignoring the common environment of full-sibs increased

heritability by 0.22 for iron and 0.28 for zinc in G_1. Improving heritability estimates may require the dispersal of family groups which will also help to remove confounding between common environment and dominance effects (random allocation of individuals rather than families), more replication (as a trade-off with selection intensity), as well as a greater understanding and control of the non-genetic factors, e.g. soil heterogeneity within a trial. Local field variation on a scale that may not be accounted for in the trial design by blocking would be perhaps better accommodated by also including spatial effects within the mixed model framework, e.g. Gilmour et al. (1997); Piepho et al. (2008b).

Difficulties in recovering useful genetic variation from unadapted wild relatives or landraces are a barrier for its use in many crop species. The Andean landrace potato is a valuable source of germplasm for potato breeding (Burgos et al. 2007), and an important part of the diet for rural populations in Peru living in the high Andes. Biofortification with iron and zinc is therefore likely to benefit poorer communities at risk of micronutrient malnutrition in this region (Burgos et al. 2007). While diploid landrace potatoes are not adapted outside the highland tropics, population improvement (e.g. phenotypic or genotypic recurrent selection) at the diploid level may increase micronutrient trait values and at the same time improve tuber shape. Taking advantage of variability for tuber dormancy and selecting for functional levels of unreduced gametes will enable transfer of gains obtained at the diploid level to more widely-adapted tetraploid populations and the development of varieties suited for new environments and wider scale deployment (Merideth Bonierbale, CIP, personal communication). The method to increase ploidy level to the tetraploid via the 4x–2x FDR (first division restitution) mechanism is well established in potato, e.g. Ortiz et al. (1991), but the effects on the genetic control of micronutrient traits in the 4x genetic background will need to be determined.

Genetic correlations

For the improvement of staple crops, breeding programmes will seek to simultaneously improve important micronutrients such as iron and zinc without detriment to yield and quality. In the present study, positive genetic correlations were found between iron and zinc, indicating that evaluation and selection for one will result in concomitant increase in the other. Repeated analysis using weaker priors of variance components resulted in a posterior mode of the genetic correlation between iron and zinc of 0.80 [0.40 0.91]. Datasets from several centres of the Consultative Group on International Agricultural Research (CGIAR) have demonstrated genetic variation and positive correlations between iron and zinc across

different genotypes of the range 0.44 to 0.61 for a number of crops including potato, maize, lentil wheat and yams (Gregorio 2002; Pfeiffer and McClafferty 2007b), but these studies gave no indication of the type of genetic control in these crops. In the present study, negative genetic correlations were found between dry matter and iron, zinc calcium and vitamin C (genetic correlation close to zero for vitamin C in G_1) when analysed on a dry- weight basis. In contrast, these genetic correlation estimates were positive (but small) when analysed on a fresh-weight basis. Although no data were available to investigate further, a possible explanation may be due to the greater concentration of some minerals at the surface layers of tubers, as is reported for minerals such as iron and calcium (Subramanian et al. 2011) . The higher surface area–volume ratio of small tubers or a dilution effect as tubers increase in size may result in the relationship between minerals and dry matter on a fresh- weight basis being confounded by tuber size (with smaller tubers tending to have a higher dry matter concentration than large tubers). A negative genetic correlation between minerals and dry matter content is not particularly helpful for breeders as a higher dry matter content is often associated with favourable sensory and cooking characteristics in potato. Consumer acceptance of new and improved cultivars has proved difficult when failing to match the preferred traditional types for certain characteristics. For example, breeding sweet potato (*Ipomoea batatas* L.) for high carotene concentration to combat vitamin A deficiency has encountered market resistance in Uganda where the preference is for white roots. An education programme has been necessary to increase consumer acceptance for orange roots in the region (T.zum Felde, CIP, personal communication). Relatively low levels of calcium are found in potato tubers and therefore potatoes are unlikely to provide a useful source of the macronutrient given the amount required in the human diet on a daily basis (Brown et al. 2012). However, as calcium deficiency in potato tubers is reported to be related to the increased incidence of several physiological disorders (Palta 1996), a better understanding of its genetic control and relationship with iron and zinc is likely to be of interest with regard to the possible consequences of long-term recurrent selection to enhance tuber micronutrient content. Further studies are required to understand the relationships between micronutrient content and agronomic and end-use quality characteristics in potato.

A further impediment to biofortification is recognised to be the relationship of target micronutrients with promoters and inhibitors that affect bioavailability upon consumption (Welch and Graham 2004). Pfeiffer and McClafferty (2007b) suggested that strategies to breed micronutrient-dense crops should consider indirect selection for bioavailability and reduced post-harvest and cooking losses as well as direct selection for increased

concentration. Vitamin C has been shown to act as a promoter that enhances the bioavailability of iron and zinc. From multivariate analysis in the present study, genetic correlations between iron and vitamin C, and between zinc and vitamin C were essentially zero, given the posterior modes and credible intervals estimated from these data.

Bayesian analysis of genetic parameters

The implementation of a MCMC procedure to fit the individual model in this study using the R package MCMCglmm (Hadfield 2010) demonstrates that Bayesian approaches, which can be applied to both Gaussian and non-Gaussian traits, are now more readily accessible to plant breeders. Other available software such as WinBUGS (Lunn et al. 2000) and MTGSAM (Van Tassell and Van Vleck 1996) have been used for tree and crop breeding data to estimate quantitative genetic parameters, e.g. Waldmann et al. (2008) in Scots Pine; Gonçalves-Vidigal et al. (2008) in common bean. Waldmann (2009) presented the case for WinBUGS as an evaluation tool for non-specialists but Hadfield's MCMCglmm-R is arguably more approachable for many plant breeders as it shares similar syntax with the popular plant breeding and trial evaluation software ASReml-R (Butler et al. 2009), a point noted by Apiolaza et al., (2011). That said, it should not be treated as a black box. The appropriate choice of priors is perhaps more important for the analysis of plant breeding data as small datasets are more typical than in animal or forest tree breeding studies and priors may therefore have a greater influence on the posterior distribution; in this context, poor prior choice will not be overwhelmed by the data. A cautionary approach is therefore required in this instance as priors can sometimes and unwittingly lead to incorrect inferences for the posterior modes due to the Markov Chains becoming trapped at a local maximum. In the present study, alternative priors were tested, following the recommendations of Gelman (2006), using the JAGS program (Plummer 2003) within R (package 'rjags'). This included a uniform prior on the variance, standard deviation and heritability as non-informative priors. Although not shown, results compared favourably with those obtained using the Inverse Gamma distribution with small equal parameters as the prior distribution for the variances which are, by default, those used in the MCMCglmm package. It seems reasonable that estimates from a previous generation (G_{n-1}) should be an appropriate choice of priors for the following generation (G_n), which was the approach taken in the present study. Blasco (2001) and Waldmann and Ericsson (2006) reviewed the advantages and disadvantages of REML and Bayesian based methods, as well as the choice of priors, when applied to the individual animal model.

2.6 Conclusions

Additive genetic effects were important for the micronutrient traits examined in this study with no detection of significant non-additive effects. Genetic correlations between iron and zinc were strong and positive. An improvement strategy employing recurrent cycles of selection may therefore optimise genetic gains in this population for micronutrients iron and zinc that are important targets for the biofortification of potato tubers. The genetic correlations between micronutrients (iron, zinc, calcium) and vitamin C were close to zero, and genetic correlations between micronutrients (iron, zinc, calcium) and tuber dry matter, an important sensory and processing character, were negative when analysed on a dry- weight basis and small but positive when analysed on a fresh-weight basis. Trial design to remove the common environment of siblings and to better account for potential local-scale field heterogeneity of mineral availability should be considered. With publically-available software such as MCMCglmm for R (Hadfield 2010), Bayesian procedures to fit the individual model are now more accessible for plant breeders to estimate variance components and genetic parameters. As well as breeding issues, it is generally acknowledged that the success of biofortification in potato and other crop species will also depend on non-genetic factors such as mineral bioavailability, palatability and the acceptance of new cultivars over traditional types.

3 Empirical and simulation studies of partially replicated (p-rep) field trials for early-stage selection in a potato breeding programme

3.1 Summary

Field data and simulation were used to investigate partially replicated (p-rep) trials as an approach to improve the efficiency of early-stage selection in a New Zealand potato breeding programme. Analysis of trial data, based on four-plant (clonal) plots planted in a p-rep design, using linear mixed models, obtained genetic and environmental components of variation for a number of yield and tuber components. Heritabilities, trial-to-trial genetic correlations and the performance repeatability of p-rep clonal selections between trial stages were high and selection was effective for economically important traits that included marketable yield, tuber dry matter content and fry quality. Simulations using a parameter-based approach, pertaining to the variance components estimated from the p-rep field trials, and the parametric bootstrapping of historic empirical data showed that improved rates of genetic gain with p-rep testing over one and two locations could be obtained compared with testing in fully replicated trials. These results indicate that the evaluation and selection of potato in fully replicated trials at the early stages of the New Zealand breeding programme may not be optimal.

3.2 Introduction

The initial stage of evaluation in a potato breeding programme comprises the visual assessment and phenotypic selection of single plants, or the evaluation and selection of families from formal progeny testing (Mackay 2007; Bradshaw et al. 2009). This is followed by more intense within-family selection from replicated clonal field trials. Under these schemes, there is reliance on an adequate multiplication rate to enter selected test entries into replicated field trials across multiple locations as early as possible. Potato, via the clonal propagation of tubers, has a relatively low multiplication rate that acts to increase the generation interval, delays testing across multiple locations and slows the rate of genetic progress and the time to deploy cultivars to industry. Furthermore, replication demands a compromise between selection accuracy and the intensity of selection; it will improve the accuracy of genotype estimates but, with the reasonable assumption that the total number of

available test plots or other test resources are fixed, the reduction in the number of tested candidates will affect the genetic response to selection.

Obtaining accurate differences between genotypes (or the 'precision' of genotype differences), and the size and significance of such differences, are important considerations for testing at advanced stages of a breeding programme and for regional variety trials. Accuracy is also desirable for early-stage trials, of course, but the emphasis here is on the ranking of a large number of test entries. For early-stage evaluation, genotypes are often considered as random effects in a linear mixed model and their BLUPs (best linear unbiased predictions) are shrunk towards the population mean accordingly (Robinson 1991; Smith et al. 2005; Piepho et al. 2008a). Under this evaluation framework, correlated data such as localised spatial field trends and pedigree information for all test entries can also be included to enhance the accuracy of evaluations and this can be further extended to the analysis of trials over multiple locations.

Increasing the amount of information will therefore result in genotype predictions becoming less conservative and closer to their true values, i.e. less shrinkage, but previous work has indicated that under certain circumstances (e.g. a high proportion of genetic to phenotypic variance), greater genetic gain may be achieved by relaxing the demands for selection accuracy by planting fewer replicated genotypes and screening a greater number of unreplicated genotypes (Bos 1983a; Gauch and Zobel 1996; Bos and Caligari 2008). Unreplicated trials are also of interest to breeders as they provide an opportunity to test genotypes for quantitative characters, such as yield, in the early stages of a breeding programme before there are sufficient quantities of seed or seed tubers available for planting in replicated trials. Trial designs are often made up of the unreplicated candidate genotypes and a number of replicated controls or 'checks' that are used for error control (Kempton 1984). Breeders may prefer this arrangement in the early stages of selection, as the randomization of genotypes is unnecessary and may be preferentially grouped, for example by family, for visual appraisal. The use of augmented trial designs was first proposed by Federer (1956) in which replicated controls are allocated by randomization into some form of systematic blocking arrangement, such as randomised complete block or row-column designs, and the remainder of the trial filled with unreplicated candidates. Checks are usually made up of a number of cultivar 'standards', but the use of these controls depletes the number of candidates that are available for selection and therefore reduces the intensity of selection. To realise any improvement in selection efficiency, there has to be a substantial reduction in plot error when check frequency is high, particularly when heritabilities are high (Kempton

1984; Kempton and Gleeson 1997). To avoid the loss of selection candidates to check cultivars, Cullis et al. (2006) and Smith et al. (2006) described partially replicated (p-rep) designs in which all entries are selection candidates, with a proportion of the candidates replicated and the remainder unreplicated. This can be extended to multi-environment trial (MET) evaluation where a proportion of candidate genotypes can be replicated within and a proportion replicated across trials. Recent studies have considered the design of p-rep trials (Clarke and Stefanova 2011; Williams et al. 2011).

In New Zealand, the evaluation of historic potato field (replicated) trials at early selection stages have observed high heritabilities of greater than 0.6 (as the proportion of genotypic or additive genetic variance to total phenotypic variance) for a number of yield and tuber traits, including marketable yield (Chapter 5) and tuber dry matter content (unpublished). Further, there is a desire to distribute genetic material to multiple locations for MET evaluation in the early clonal stages as soon as possible. This has motivated the exploration of partially replicated (p-rep) trials for early-stage potato selection in the New Zealand potato breeding programme to increase selection efficiency, as previous observations suggest that full replication in breeders' trials may not be optimal.

In this study, genetic parameters and the repeatability of clonal performance for a number of important yield and tuber quality traits, based on selections from p-rep trials were measured. Variance components estimated from the p-rep trials provided the basis for inference from parameter-based simulations to determine the expected responses to selection. Evaluations of simulated data were over one or two 'environments' and used a linear mixed model with varying numbers of tested genotypes, heritabilities and genetic covariances for a single stage of selection. A second method of simulation based on the assessment of historical trial data used a parametric bootstrapping approach, with trial analysis based on a formal p-rep design structure. Based on the findings of this study, the use of partial replication in early stage selection trials as a means to improve selection efficiency in a potato breeding programme is discussed.

3.3 Materials and Methods

Trial data

The genotypes tested were random selections from single-plant plots (clonal stage 1 or C1 trials) taken in March 2011 and were made up of 44 full-sib and 17 half-sib families. The selections were representative of genotypes screened as part of the potato breeding

programme at The New Zealand Institute for Plant & Food Research Limited (PFR). Field trials were planted at the Lincoln PFR research site in the South Island of New Zealand (Fig. 1-4). In the first year (2011-2012), a partially replicated (p-rep) field trial was designed using DiGGer (Coombes 2011), a Windows console program, by supplying an input file (Fig. 10-2, Appendix II). The trial was designed to be resolvable for the complement of replicated genotypes in two dimensions (i.e. across row-blocks and column-blocks (Fig. 3-1), and consisted of 236 entries with two checks ('Fraser' and 'Agria'). It was planted as four-tuber (four-by-one) plots in October 2011 and harvested in April 2012, 165 days after planting (and is hereafter referred to as $C2_1$; note that trials given the same subscript were grown in the same season). A p-rep trial was designed and grown again in the 2012-2013 season ($C2_2$) with 200 of the same entries from $C2_1$ and the same two checks. The target replication level of both C2 trials was p=1.20 (where 20% of test entries are replicated), but this was reduced to approximately p=1.18 in the first year because of attrition of genotypes. After analysis of the p-rep 2011-2012 trial, 48 genotypes were randomly selected and planted in a trial in the 2012-2013 season ($C3_2$). The trial was designed as a Latinized row-column with CycDesigN v4.0 (CycSoftware 2009) and planted with three replicates in six-tuber (six-by-one) plots. Both the $C2_2$ and $C3_2$ trials were planted in October 2012 and harvested in April 2013, 169 days after planting.

After harvest, measurements were made for a number of agronomic and tuber characteristics, including total and marketable yield, percentage marketable yield, percentage dry matter content, mean tuber weight and fry quality. Plot yield was recorded at harvest as both a total tuber yield and a marketable tuber yield for analyses. Marketable (market) tuber yield was the graded yield after undersized (less than 80 g) and defective tubers had been removed. Defective tubers, for example, may have been afflicted with secondary or abnormal growth, rot or excessive greening. Yield was also expressed as the percentage marketable fraction of the total yield and is referred to as the percent market yield. These data were logit transformed so that, $ln\left(\frac{p}{1-p}\right)$ where p is the proportion of market to total tuber yield, if residual plots did not meet assumptions of normality. For each (marketable) yield measurement, a mean tuber weight (g) was calculated and dry matter content was measured using a weight-in-air/weight-in-water conversion calculation. Four tubers from each plot were also sampled,

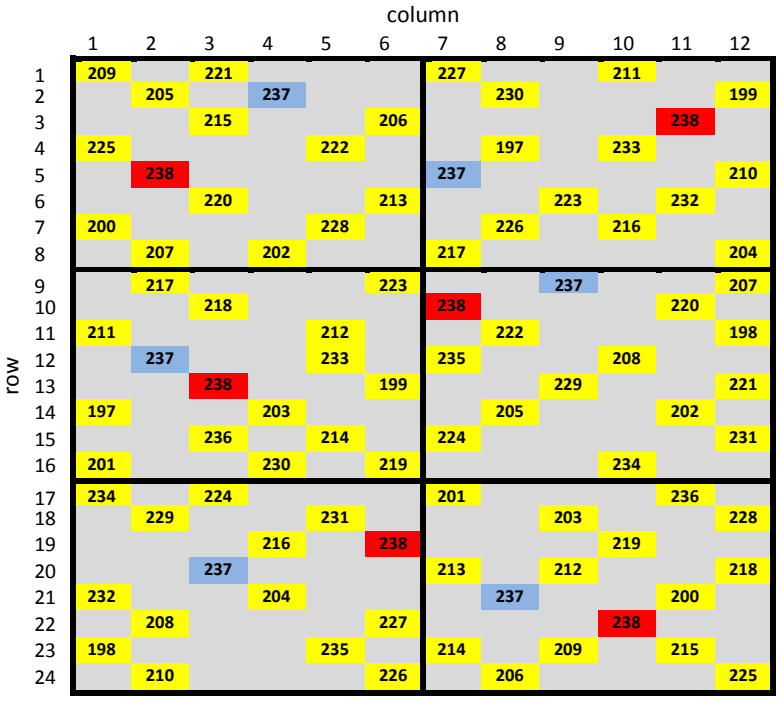

Figure 3–1 Partially replicated (p-rep) potato trial with 1:196 unreplicated entries (not labelled), 197:236 replicated entries and two checks (237:238)

chemically treated with sprout suppressant and stored at 10°C. These tubers were removed from cold storage in September after 120 days and held for 24 h under ambient conditions. Potato slices (crisps) were then cooked by frying for 2.5 min at 190°C in canola oil and scored on a 1 to 9 scale (Fig. 3-2) for fry score, with 1 indicating a high fry quality (light coloured, with no evidence of discolouration) and 9 indicating a very poor fry quality (blackened discolouration).

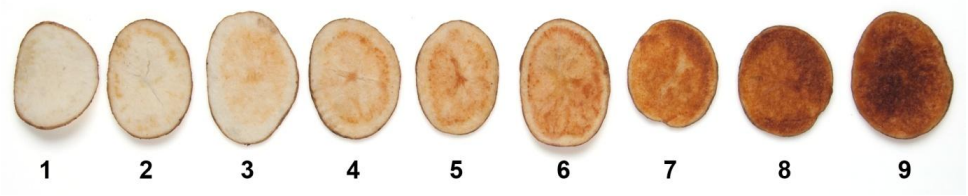

Figure 3–2 Fry score potato colour assessment scale where 1=very high quality (no discolouration) and 9=very poor quality (blackened discolouration)

Statistical analysis

Trials were analysed with the general form of a univariate linear mixed model, $\mathbf{y} = \mathbf{X}\boldsymbol{t} + \mathbf{Z}\boldsymbol{u} + \mathbf{e}$ where **y** is the vector of trait observations, ***t*** is a vector of fixed effects, **u** is a vector of random effects which, in this instance, included subvectors of random (non-genetic) design factors, $\boldsymbol{u_b} \sim N(0, \mathbf{I}\sigma^2_b)$, e.g. replicate and/or block (within replicate) and $\boldsymbol{u_g} \sim N(0, \mathbf{I}\sigma^2_g)$ as subvectors of random genetic effects respectively. The vector of random error terms is given by $\boldsymbol{e} \sim N(0, \mathbf{I}\sigma^2_e)$ while **X** and **Z** are known incidence matrices for the fixed and random effects respectively, **I** are identity matrices and the subscripted σ^2 is the variance of each of the random effects.. For each trait, analyses considered the blocking structure of the trial designs, allowing for independent random effects, such as blocks, and independent plot errors. For the $C2_1$ and $C2_2$ trials, check cultivars were fitted as fixed effects. There were no checks planted in trial $C3_2$. A fixed covariate was fitted in $C2_1$ and $C2_2$ to account for some waterlogging that had occurred for a short period over the duration of the trials. A fixed covariate was also fitted to account for the loss of plants in a small number of plots, but any plot with two plants or fewer was considered as a missing value. For tuber yield traits and dry matter content, a spatial model was tested by separating the random error term **e** into spatially dependent (autocorrelated) and spatially independent errors, following the AR1 spatial correlation model of Gilmour et al. (1997). A likelihood-ratio test was used as the criterion to test for the importance of the fitted spatial effects, which remained as an addition to the blocking features of the trial if model fit was improved. The main emphasis was the selection of candidate varieties based on their genotypic values to enter the next stage of clonal selection, but the selection on breeding values was also considered. Therefore, data were also analysed after replacing the independent genotypic variance given by $\mathbf{I}\sigma^2_g$ with a pedigree-based genotypic variance given by $\mathbf{A}\sigma^2_a$, the variance-covariance matrix of the additive genetic effects (breeding values), where **A** is the numerator relationship matrix that provides the between-genotype relationship as two times the coefficient of coancestry. .The pedigree was built from PFR field books and an online potato pedigree database (van Berloo et al. 2007). Variance component estimates provided an indication of the magnitude of signal-to-noise expected from p-rep potato trials and were used as basis to infer the expected response to selection using parameter-based simulations (see p. 50). In general, heritabilities were obtained from either the proportion of additive to phenotypic variance (h^2), or the proportion of genotypic to phenotypic variance (H^2, excluding the pedigree), with the phenotypic variance including the genetic, block/replicate and error variances.

As variance components are unknown, the empirical genotypic and breeding values (EBVs) were obtained from the BLUPs (best linear unbiased predictors) of random effects (e.g. Smith et al. 2005, p.458). Coefficients of correlation were obtained for the performance of common genotypes between C2 and C3 trials based on their BLUPs of genotypic and breeding values. The univariate model was then extended to a bivariate analysis in order to estimate genetic variances and covariances simultaneously for the same trait measured in different environments (synonymous with trials), e.g. between $C2_1$ and $C3_2$, and to estimate a genetic correlation for each trait.

Simulation of genetic response: parameter-based simulation

The first approach to simulation used a stochastic, parameter-based method to model selection in replicated and p-rep trials over both one and two locations for a single clonal selection stage. In predicting the response to selection, it was assumed that the trait under consideration was normally distributed. Normal distributions of true genotypic values (g) and environmental deviations were obtained from given estimates of genetic and environmental variances, and these were used to produce a simulated breeding population. In this case, the vector of additive effects a and non-additive effects d were assumed to be mutually independent, so that the vector of total genetic effects ($g = a + d$) had distribution $g \sim N(0, \sigma_a^2 I + \sigma_d^2 I)$. Alternatively, the additive and non-additive effects could be sampled independently from separate distributions when assuming no covariance between additive and non-additive genetic effects. In the present study, the difference between the two sampling strategies is likely to be small, but independent sampling allows for greater flexibility, e.g. allowing for additive genetic covariances between related individuals, if a pedigree structure is incorporated, and including dominance as a separate component. For p-rep trials, the level of replication was set at 25% (p=1.25), which was compared with selection from trials with two replicates (p=2) Assuming that the genetic values for the trait were polygenic, the phenotypic variance was arbitrarily set at 10, with heritability varying from 0.1 to 0.8, giving an equivalent signal-to-noise ratio (σ_g^2 / σ_e^2) that ranged from 0.11 to 4.0 (Table 3-1). For selection over two locations, the sampling of genetic values was from a multivariate normal distribution with the same genetic variances and genetic correlations of 0.2, 0.5 and 0.8. The sites were assumed to be equally weighted so that the true genotypic value for each individual was the mean of the sampled genotypic values. Three test scenarios were simulated for each of p=1.25 and p=2:

Table 3–1 Parameter values applied to the simulation of replicated and p-rep trials

Phenotypic variance (σ_p^2)	Genetic variance (σ_g^2)	Environmental variance (σ_e^2)	Heritability	Ratio: (σ_g^2/σ_e^2)
10	1	9	0.1	0.11
10	2	8	0.2	0.25
10	3	7	0.3	0.43
10	4	6	0.4	0.67
10	5	5	0.5	1.00
10	6	4	0.6	1.50
10	7	3	0.7	2.33
10	8	2	0.8	4.00

A: single location testing, fixed number of total plots, nP (total no. of plots) = 100, (to correspond approximately with simulation using historical field data), and 1000.

B: extension of (A) with testing over two location for both p=1.25 and p=2, fixed number of total plots, nP = 2000, distributed over two locations.

C: Fixed number of total plots, nP = 2000, distributed over one (p=2) or two (p=1.25) locations. The replication level at p=1.25 to test 800 genotype entries (1000 plots in each of two locations) is shown in Table 3-2. Testing at p=2 at one location (2000 plots) allows 1000 genotypes in total to be tested.

Table 3–2 Replication level for p=1.25 over two locations, fixed number of total plots = 2000 testing 800 genotypes

Entry no.	Location 1	Location 2	Total replicates
1:200	2	1	3
201:600	1	1	2
601:800	1	2	3
Plots	1000	1000	

For each scenario, data were generated for 10,000 simulations and analysed using a linear mixed model in ASReml-R (Butler 2006) to obtain predictions of the empirical genotypic values. A relative genetic response to selection ($\Delta G'$) for p-rep tested genotypes (relative to the response when p=2, so that R_P/R (where R_P is the p-rep selection response and R is the replicated selection response) was also calculated and stored for each simulation run; the top performing individuals from the p-rep analysis, comprising 5, 10 or 20% of the total genotypes tested (*s*), were selected based on the ranking of their empirical genotype values. The selection response was considered to be the difference between the mean true genotypic values of the *s* individuals and the mean of the true genotypic values of the breeding

population. Therefore, truncation selection over a single cycle on a single trait was applied and the selection intensity was obtained directly from the proportion of genotypes selected when p=1.25 for scenarios A and B and p=2 for scenario C.

Simulation of genetic response: bootstrap resampling using historic empirical data

The second approach was an empirically based simulation that aimed to take into account the error structure from historical field data and overlay a formal p-rep design on the original replicated trial. The field data were based on Pukekohe early-stage potato yield trials in the North Island of New Zealand, which are small, multiple α-Latinized designs (80 to 100 genotypes per trial) of two replicates with plots made up of 12 clones grown in a six-by-two arrangement. The selection stage for these trials was equivalent to the C2 stage at the Lincoln site. The bootstrap simulation is outlined is described as follows:

1. The replicated trial was analysed using a linear mixed model in ASReml-R (Butler 2006), following the general form of the univariate model outlined in the section *Statistical analysis*. Genotypes were considered to be random effects and the overall mean fitted as a fixed effect. The residuals from this analysis were used to give the spatial layout of the environmental effects for subsequent p-rep analyses for each trial. The best linear unbiased predictions of genotypic values for tested genotypes were obtained from the solutions of the mixed model equations using the estimated variance parameters. The resulting empirical genetic values (eGVs) were considered to be the actual genetic values (aGVs) for the simulations.

Table 3–3 Example of the expected number of total genotypes, replicated genotypes and total plots available for p-rep simulation at different levels of p-rep (p) using the empirical data from an historic replicated trial (step 2, p. 52) with a total of 160 plots and 80 genotypes. At a minimum p of 1.15, the total number of plots available for simulation is 92. Higher levels of p therefore require a random elimination of genotypes for each set of simulated data

	p (level of p-rep)							
	1.15	1.25	1.35	1.45	1.55	1.65	1.75	2.00
Total genotypes	80	74	68	63	59	56	53	46
Replicated genotypes	12	18	24	29	33	36	39	46
Total plots	92	92	92	92	92	92	92	92

2. A minimal level of p-rep, P_{min}, was elected to be 1.15, and therefore the standard p-rep trial size was $N_g \times P_{min}$ where N_g = number of genotypes in the replicated trial. For example, a replicated C2 trial of 160 plots would comprise 80 genotypes and so a p-rep analysis of this

particular trial would be 80 × 1.15 = 92 plots in total for all levels of partial replication. Levels of p-rep (p) greater than 1.15 therefore required the random elimination of some genotypes for each analysis, as increasing replication would not allow the full complement of genotypes to be assessed at any one time over a fixed number of plots (see Table 3-3).

3. For each level of p-rep, p, a trial was designed using DiGGer (Coombes 2011) in R (R Development Core Team 2012). The trial was designed to be resolvable, for the complement of replicated genotypes, in two dimensions (i.e. across row-blocks and column-blocks). This design was randomly located over the replicated C2 trial with associated plot residuals (as initially computed in step 1) allocated to the new layout. Genotypes were then randomly allocated to the treatment numbers of the trial design created. For computing expedience, designs were used several times by transformation.

4. For each plot, a simulated genetic value was generated from the parametric bootstrap of the genetic value (pbGV) for the genotype plus the plot residual (environmental) effect for its location in the trial (Fig. 3-3). The pbGV was obtained by adding random noise taken from a normal distribution with mean = 0 and standard deviation = se to the aGV, where se is the standard error of the eGV as obtained from the replicated C2 analysis in 1.

5. The data generated were then analysed in ASReml-R (Butler 2006) following the general form of the univariate linear mixed model as described in section *Statistical analysis*. A relative response to selection, $\Delta G'$, was also calculated and stored for each simulation run p

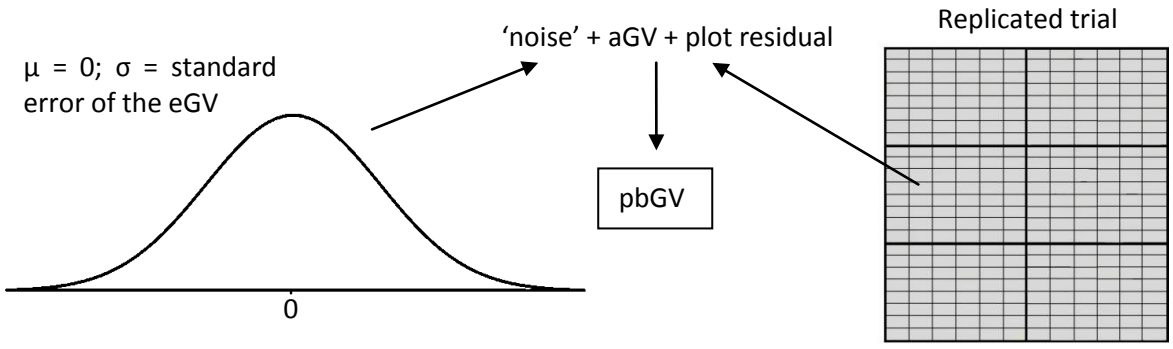

Figure 3–3 A simulated genetic value (pbGV) for a genotype in each plot was generated (step 4, p. 53) by adding the plot residual from its particular location in the trial and the random noise, taken from a normal distribution with mean = 0 and the standard deviation = the standard error of the empirical genetic value (eGV), to the actual genetic value (aGV). The eGVs (= aGVs) and plot residuals were obtained from the replicated trial analysis (step 1, p. 52)

following the '*parameter-based simulation*' approach as outlined, a method previously described by Piepho and Möhring (2007). Only marketable yield, as the character of most interest, was tested. There were a total of 5000 simulation runs for each p.

3.4 Results

Field trials

Variance components and broad-sense heritability estimates for the p-rep trials $C2_1$ and $C2_2$ are presented in Table 3-4. Estimates of narrow-sense heritabilities were very similar (results not shown). Heritabilities for dry matter were particularly high, with replicate and block effects close to the boundary, and were subsequently dropped from the model for both $C2_1$ and $C2_2$ (and $C3_2$). Spatial correlations of 0.42 and 0.54 (rows) were obtained for dry matter at C2. Spatial correlations were much higher for $C2_2$ than for $C2_1$ for total and marketable yields. This may have been due to soil compaction which was observed in some areas of the trial, causing spatial patchiness because of periods of waterlogging (or as a result of poorer root development because of soil panning) during the growing season. A significant fixed covariate ($P<0.05$) fitted for the analysis of tuber yield traits in $C2_2$ to account for the worst affected area disappeared when spatial effects were fitted. This contrasts with $C2_1$, in which both a fixed covariate ($P<0.05$) and spatial effects were fitted. The waterlogging in this trial was only found in rows one to two (Fig. 3-1) but there was a greater weed burden throughout the trial which also may have contributed to the spatial heterogeneity.

Correlations of genotypic values estimated from univariate analysis for $C2_1$ and $C3_2$ (between adjacent seasons), which is the comparison of most interest, were high for all traits considered, with the exception of percent marketable yield (Table 3-5). These were very similar to the correlations between EBVs, which are displayed graphically in Figure 3-4. Correlations between C2 and C3 grown in the same year ($C2_2$, $C3_2$) were higher than or at least as high as correlations between C2 and C3 grown in different years ($C2_1$, $C3_2$). In general, the differences were not that great, however (with the exception of percent marketable yield and fry score). Differences may have been due to both seasonal effects and carry-over (or 'maternal') effects (for yield components) from growing test plots with tubers selected from single plants for each genotype. Figure 3-4 illustrates the strong relationships identified for total yield, marketable yield and dry matter EBVs between stages $C2_1$ and $C3_2$, with a weaker relationship evident for percent marketable yield.

Correlation of genotypic values between $C2_1$ and $C3_2$ were usually higher than those between the correlation of observed phenotypic values in $C2_1$ and $C3_2$ (Table 3-5). This difference was often small (total yield, fry score), with no difference found for dry matter, illustrating that the accuracy of phenotypic assessment was comparable with that given by genotype evaluation. The greatest difference was for marketable and percent marketable

yield. Trial-to-trial genetic correlations between $C2_1$ and $C3_2$, estimated directly from bivariate analyses, were all high (although mean tuber weight did not converge), particularly for yield (total and marketable) and dry matter. These estimates were similar but slightly lower than those between $C2_2$ and $C3_2$.

Table 3–4 Summary of variance component and heritability (H^2) estimates from p-rep potato trials $C2_1$ (upper values) and $C2_2$ (lower values)

Trait	σ_g^2	σ_r^2	σ_{rb}^2	ρ_{row}	ρ_{col}	σ_e^2	H^2 (se)
Total tuber yield (kg)	8.9	0.1	0.2	0.28	-	1.2	0.86 (0.04)
	4.9	-	-	0.79	0.72	1.6	0.76 (0.06)
Marketable tuber yield (kg)	7.8	-	0.2	0.30	-	1.4	0.83 (0.04)
	4.5	-	-	0.76	0.65	1.9	0.70 (0.06)
Percent marketable tuber yield	0.47	-	0.03	0.17	-	0.25	0.65 (0.07)
	0.38	0.01	-	0.37		0.22	0.62 (0.07)
Dry matter (%)	2.82	-	-	0.42	-	0.34	0.89 (0.02)
	3.32	-	0.06	0.54	-	0.69	0.82 (0.04)
Mean tuber weight (g)	2887	-	152	-	-0.18	1414	0.65 (0.07)
	1116	-	-	0.22	0.31	397	0.74 (0.05)
Fry score	2.4	0.05	-	-	-	0.58	0.79 (0.05)
	2.73	-	0.03	-	-	0.68	0.79 (0.05)

σ_g^2 is the genetic variance; σ_r^2 *and* σ_{rb}^2 are the replicate and block/replicate variances, ρ_{row} and ρ_{col} are the spatial correlation parameters; σ_e^2 is the residual error variance; H^2 is the estimate of the heritability – the proportion of genetic variance to the phenotypic variance, and se is the standard error of the heritability estimate from Taylor series expansion (Lynch and Walsh 1998)

Parameter and empirical-based simulations

At a selection proportion (*s*) of 5% with 100 tested genotypes at a single location (scenario A), parametric-based simulation showed that the relative selection response was close to unity when h^2 (or H^2) ≈ 0.4 – 0.5 (Fig. 3-5A1). The relative response reduced slightly at all levels of *s* (5, 10 and 20%) with 1000 tested genotypes (Fig. 3-5A2). There was some evidence to suggest that relative gain was overestimated with small sample sizes and at low heritabilities, possibly because of difficulties in estimating variance components. At *s*=20, there was no advantage in replicating trials. When testing over two locations with both full

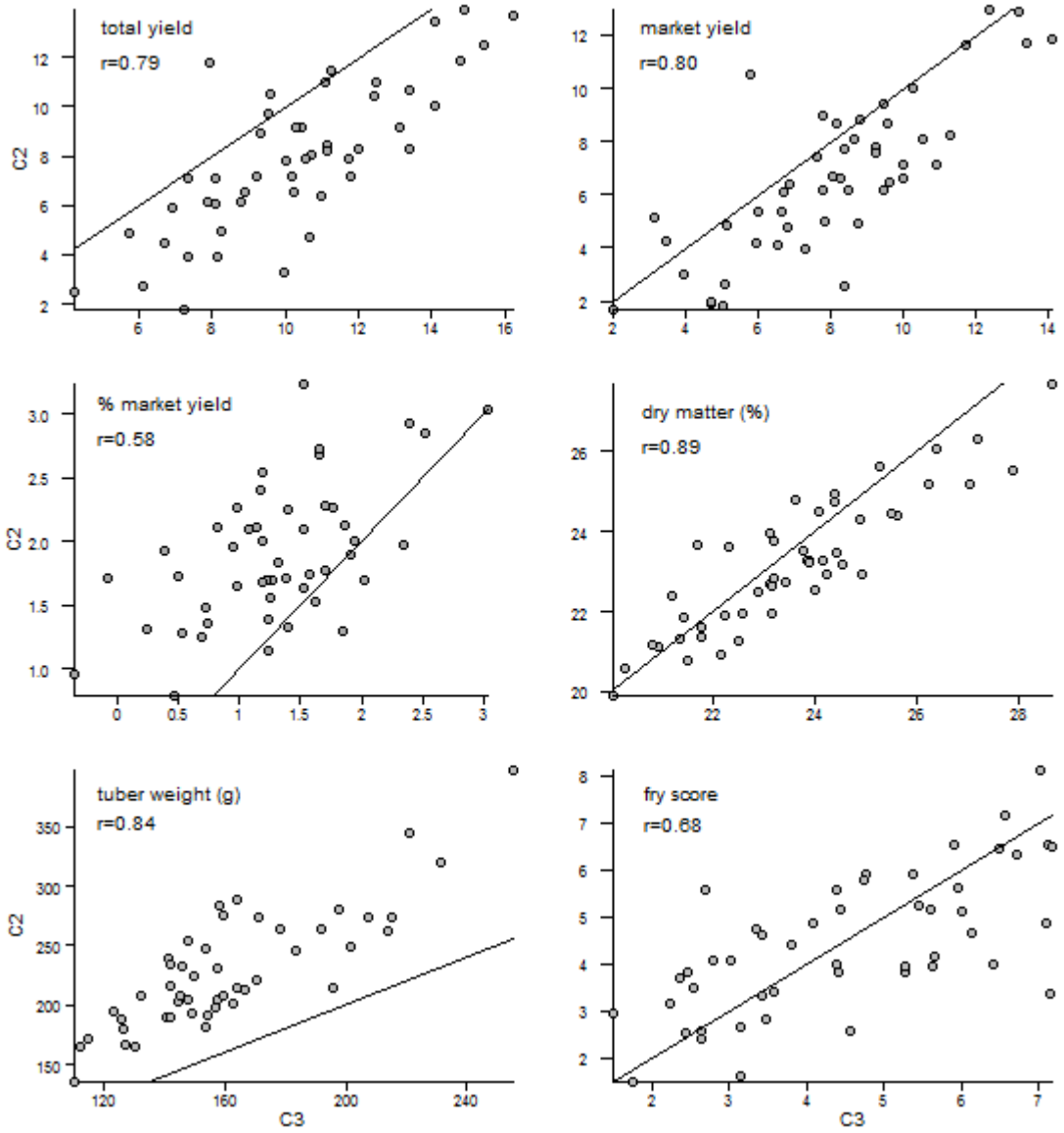

Figure 3–4 Correlation of empirical breeding values (EBVs) between $C2_1$ and $C3_2$ for potato yield and tuber characteristics, showing the line of unity

replication and p-rep (scenario B) at a selection proportion of 5%, relative response was at unity at a heritability of just over 0.30 when the trial-to-trial genetic correlation was high (0.80) and just over 0.40 when the genetic correlation was low (Fig. 3-5B). This reduction in heritability when $R_p/R=1$ compared with scenario A is expected, given that some genotypes were replicated three times (four times when p=2) over the two locations. The relative responses to selection for all correlations did not surpass 1.10 and tended to converge as heritability approached 0.80. The advantages of extending p-rep testing to two locations over fully replicated testing at a single location on a fixed number of plots, particularly for low

trial-to-trial correlations, are clearly seen in Figure 3-5C. Relative response easily offsets the reduction in the number of p-rep genotypes tested, even at very low heritabilities. The advantage of testing over two locations when trial-to-trial correlations are high is generally small, however, with the relative response trending towards unity as heritability increases, although R_p/R is greater than 1 for all heritabilities tested.

Table 3–5 Correlation coefficients of genotypic (and phenotypic) values of potato yield and tuber characters; C2 (p-rep) trials and C3 (replicated) trials

	[†]r			[‡]r_g
Trait	$C2_1,C3_2$	$C2_2,C3_2$	[§]obs_1,obs_2	$C2_1,C3_2$
Total tuber yield (kg)	0.79	0.84	0.72	0.89
Marketable tuber yield (kg)	0.78	0.78	0.71	0.89
Percent marketable tuber yield	0.48	0.71	0.41	0.75
Dry matter (%)	0.89	0.93	0.88	0.95
Mean tuber weight (g)	0.77	0.81	0.77	[¶]nc
Fry score	0.70	0.82	0.67	0.76

[†]Correlation of genotypic values from univariate analyses or, [§]correlation of observed (obs) phenotypic values between $C2_1$ and $C3_2$; [‡]Genetic correlation estimated from bivariate analysis. [¶]nc = non-convergence

Table 3-6 shows an example of the performance of a bootstrap simulation run (5000 samples) at all levels of tested p-rep using historical field data, with a target number of test plots set at 92. The total number of genotypes tested at p=1.25 was approximately 75, which reduced to 46 at p=2. Trials presented in Table 3-7 are a representative set of trials, with regard to heritabilities, from 1999 to 2012 for marketable yield, with the lowest and highest heritabilities found in trials C2-06A and C2-99D/C2-00B respectively. For trial C2-12E, H^2 = 0.46 and the relative response was close to one at s=0.05, which was similar to the result from parametric-based simulation (Fig. 3-5A1). For trial C2-06B, H^2 = 0.66 and the relative response was 1.35 at s=0.20, which was inflated well above that which was expected, given the results of the parametric-based simulation at this heritability (Fig. 3-5A1). There were few Pukekohe trials with heritabilities of less than 0.4 for any yield traits with records available, and so there was limited opportunity to test the empirical simulations at low heritabilities. An exception was the C2-06A trial, where the heritability of marketable yield in this trial was estimated to be 0.25. At p-rep=1.25, simulation of this trial gave a relative selection response of 0.93 for s=5 and 1.05 for s=10. From parametric-based simulation results, unity of relative selection response (Rp/R=1) for s=10 was expected at an

approximate heritability of 0.3. Again, the relative selection response appeared to be inflated above expectation, when s=20.

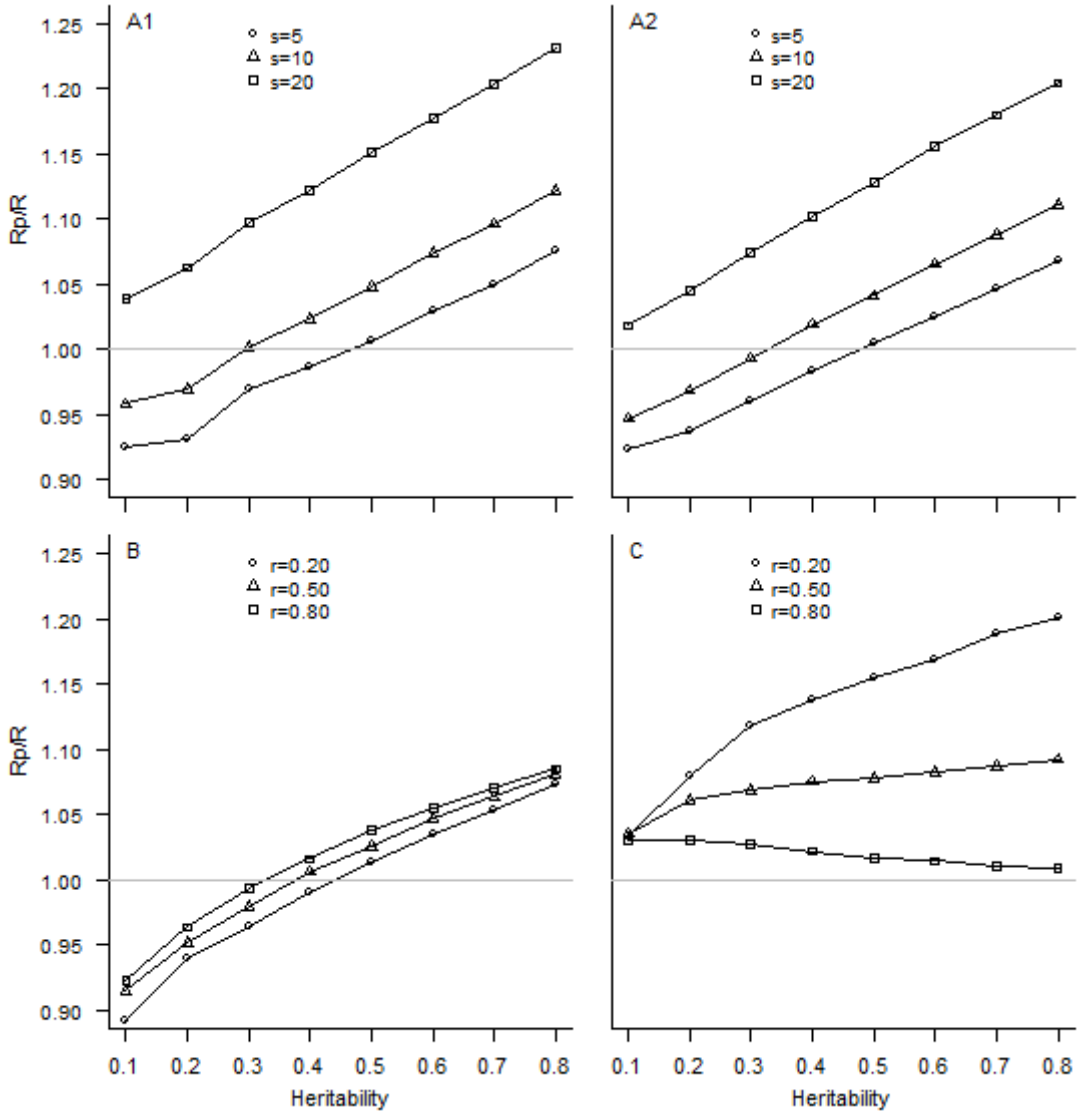

Figure 3–5 Simulation of the relative response to selection (R_p/R) at increasing heritabilities where A1: testing in one location, 100 total plots and 5, 10 and 20% proportion selected (s); A2: as A1 but with 1000 total plots; B: testing in two locations with 2000 total plots (1000 in each location) with genetic correlations (r) between locations of 0.2, 0.5 and 0.8 and s =5; C: testing in one location (replicated), 1000 genotypes, 2000 total plots and s = 5, or testing in two locations (p-rep), 800 genotypes and 2000 total plots (1000 in each location), s=5 and genetic correlations as in B

Table 3–6 Summary of the means from the analysis of Pukekohe C2 potato trial for the p-rep bootstrap simulation of marketable yield, as an example. The replicated trial consisted of 160 plots in total and 80 genotypes. The target number of plots for the simulation in this particular example was 92. Actual p is the level of p-rep obtained in the simulation

	target p (level of p-rep)							
	1.15	1.25	1.35	1.45	1.55	1.65	1.75	2.00
actual p	1.15	1.24	1.35	1.47	1.57	1.64	1.76	2.00
total genotypes	78.7	74.5	69.0	64.1	59.6	56.0	53.5	46.2
replicated genotypes	12.1	18.3	24.5	29.9	33.8	36.1	40.5	46.2
plots	93.0	92.7	93.6	93.9	93.4	92.1	93.9	92.4
rows	16.0	16.2	16.2	16.3	16.1	15.0	14.7	11.6
columns	5.9	5.9	5.9	5.9	6.0	6.3	6.5	8.0

Table 3–7 Summary of the p-rep bootstrap simulation using 1999 – 2012 empirical data (Pukekohe early stage trials) for potato marketable yield (t ha^{-1}) at p-rep=1.25

Trial	N_g	H^2	ΔG (5)	ΔG′ (5)	ΔG (10)	ΔG′ (10)	ΔG (20)	ΔG′(20)
C2-99B	80	0.56	10.02	1.01	8.53	1.08	6.81	1.21
C2-99D	80	0.71	28.01	1.16	21.75	1.22	16.08	1.34
C2-00A	80	0.69	18.37	1.04	16.03	1.11	13.00	1.23
C2-00B	80	0.71	20.54	1.02	16.96	1.08	13.37	1.20
C2-06A	90	0.25	2.81	0.93	2.66	1.05	2.26	1.17
C2-06B	90	0.66	15.87	1.07	13.93	1.17	11.18	1.35
C2-07B	80	0.57	6.68	1.02	5.75	1.08	4.52	1.23
C2-07C	80	0.64	8.47	1.02	7.58	1.07	6.29	1.20
C2-12D	98	0.57	13.70	1.04	11.68	1.10	9.12	1.24
C2-12E	91	0.46	7.76	1.00	6.88	1.04	5.60	1.15

N_g = number of tested genotypes at p=1.25; H^2 is the heritability of the replicated trial; ΔG is the genetic gain and ΔG′ is the relative genetic gain.

3.5 Discussion

Reducing the degree of clonal replication in the present context is motivated by two main factors: the relatively low multiplication rate of potato tubers, i.e. a shortage of planting material at the early stages and therefore the time lag associated with entering candidates into formal trial evaluation and; a desire to increase the number of candidates tested (when the total number of plots is fixed and it is assumed that the number of extra candidates that will need to be phenotyped is not restricted). The results are presented in terms of genetic gain, but it is acknowledged that other factors will need to be considered under a broader context. These may include the net worth of genetic gain, the biological constraints of the crop, the physical and financial restrictions that are imposed upon the programme and the opportunity costs, particularly when extending trials beyond a single test location. These issues are

beyond the scope of the present study, and the motives and constraints of implementing a selection scheme will obviously vary from programme to programme.

Empirical and simulated data

For the tuber yield and quality traits presented, the results indicate that increased genetic gain could be achieved in the PFR potato breeding programme by applying p-rep trial evaluation at the early stages of selection. Given the heritability levels reported, previous work based on order statistics and known selection formulae also indicate that the expected genetic gain at moderate to high heritabilities may be greater at a single location by planting fewer replicated entries (Bos 1983a; Gauch and Zobel 1996). This moderates the selection accuracy but the trade-off is that it allows a greater number of genotypes to be tested (Bos 1983a). From Figures 3-5A1 and 3-5A2, the relative selection response was close to unity when h^2 (or H^2) ≈ 0.4 – 0.5, which was similar to that predicted by Bos (1983a; Table 3) for unreplicated testing that assumed an infinite number of plots and genotypes. The recommendations of Gauch & Zobel (1996) were more conservative, so that unreplicated testing was suitable when h^2 = 0.6 and greater, when the total number of plots was 100. At 1000 plots, two replicates were favoured when h^2 was 0.75, the maximum shown. This work accounted for the extra noise generated by testing more plots, so that increasing the number of genotypes tested had the desirable effect of increasing the number of superior genotypes, but also the undesirable consequence of adding more noise (inferior genotypes). There was only a small indication of this trade-off in the present study, as the relative response reduced slightly at all levels of *s* with 1000 test plots (Fig. 3-5A2) compared with 100 test plots (Fig. 3-5A1). The deterministic approach of Bos (1983a) to calculate genetic gain at different levels of replication, with the ratio of the relative response when the number of total plots and selected candidates is fixed, was presented as $\frac{R_p}{R_k} = \frac{k}{i}\sqrt{\frac{1+(k-1)h^2}{k}}$ where p = 1 (non-replicated), i is the selection intensity at p = 1 and k is the number of replicates (≥ 2). Cullis et al. (2006) and Piepho and Möhring (2007) have emphasised that under more complex analysis, for example, when data is unbalanced and genotypic effects are correlated, it is not appropriate to apply the standard measures of heritability obtained to compute a response to selection. Other measures of heritability under such situations have therefore been proposed, including a 'generalised' heritability of Cullis et al. (2006), given as $h^2 = 1 - \frac{mean\,(PEV)}{\sigma_g^2}$ where σ_g^2 is the genetic variance estimate and PEV is the vector of estimated prediction error variances (the measure of variation about the predicted genotype value) for the random genotype empirical BLUPs.

These are obtained from the inverse of the coefficient matrix of the mixed model equations. Piepho and Möhring (2007) proposed an alternative approach by measuring selection response directly using a simulation-based method, thus avoiding the necessity to define the heritability or use an inappropriate measure of heritability altogether. This simulation approach has been used in the present study, as well as Chapters 4 and 5.

Block and residual error variances were low for dry matter content in all trials (C2 and C3), which showed a particularly high heritability. Using a specific gravity measurement (from the air and underwater weight of tuber samples) to estimate dry matter is labour intensive and time consuming and the current practice of taking a composite sample (by pooling replicate plots) of tubers for testing is therefore supported. Methods of sampling and analysing composite samples based on partial replication, for traits that are time-consuming or costly to measure, have been explored by Smith et al. (2006; 2011).

Performance repeatabilities of genetic values between stages were mostly high and results from parameter-based simulations also found an expected increase in the response to selection at the selection intensities that are typically employed in the early stages of a clonal selection programme. In Scotland, work by Caligari et al. (1986) showed that the repeatabilities in the phenotypic performance of potato in early-stage clonal trials for total tuber yield, mean tuber weight and the number of tubers were poor, indicating that evaluation of the phenotype was a poor predictor of genotype performance in subsequent years. This was particularly apparent for the relationship between the first clonal stage and second clonal stage but less so between the second and third clonal stages, which were equivalent to the C2 and C3 stages of the present study. Correlation coefficients of 0.52 and 0.60 between the second and third clonal stages were obtained for total tuber yield and mean tuber weight (Caligari et al. 1986; Table 2) over two locations, although there was no significant clone × location interaction. This compared with respective estimates for the same traits of 0.79 and 0.77 (genotypic values) and 0.72 and 0.77 (phenotypic values) (Table 3-5). Although the approaches in analyzing the data were different, the reasons behind the differences in correlation coefficients found between the second and third clonal stages in the two studies may serve to demonstrate that the results presented here cannot necessarily be extended to other programmes because of differences in genetic parameter estimates between breeding populations and the vagaries in the accuracies of estimating genotype effects over different seasons and locations.

Early-stage testing in multiple locations

The results from simulation also showed that selection over two locations in p-rep trials may be particularly beneficial compared with selection over one location in a replicated trial when locations are weighted equally in a selection index, i.e. selection is for broad adaptation. Positive correlations between locations only were considered in this study, as negative correlations have not been found in previous analysis of tuber yield traits from MET data collected over a number of years from the PFR breeding scheme. Negative correlations indicate a greater importance of qualitative G×E interaction and suggest that alternative approaches to selection may be required, i.e. separate selection schemes to target specific adaptation in subdivided regions, e.g. Atlin and Frey (1990); Atlin et al. (2000); Windhausen et al. (2012).

For METs, the precision of across-trial comparisons is compromised in the presence of genotype × environment (G×E) effects. The magnitude of estimated average genotype × environment variances for yield in a number of different crops has previously been reported (Talbot 1984). For potato, this was found to be to be greater than the within-trial plot variances, and so efforts expended on maximising selection precision by replication on an individual trial site may therefore be wasted (Kempton, 1984; (Kempton and Gleeson 1997). At low heritabilities, the advantage of extending p-rep testing to two locations is maintained over full replication at a single location (scenario C), particularly for low trial-to-trial correlations (Fig. 3-5C) and despite a reduction in the number of candidates tested on a fixed number of plots. This is a more realistic scenario than B at the C2 stage of selection, because of the shortage of available planting material. The advantage of testing over two locations when trial-to-trial genetic correlations are high is not so obvious, however, with the relative response trending towards unity as heritability increased. Managing trials over two or more locations may become more difficult to justify in this case, with the difference in gain (which was less than 1%) having to be measured against the extra costs incurred, but this is beyond the scope of the present study. Unreplicated or p-rep trials offer a means to increase the number of test genotypes over a fixed number of plots or, alternatively, a means to reduce the number of plots to test an equivalent number of genotypes which may maximise gain per unit cost (Stendal and Casler 2006). Moreau (2000) found that unreplicated trials were optimal for both phenotypic selection and marker-assisted selection when traits were sensitive to G×E effects. For marketable yield, previous analyses of PFR early stage trials in Pukekohe have shown that heritabilities are generally high and that genetic correlations between adjacent

seasons are also usually high (>0.70) (Chapter 5). This may not be the case for all traits, however as, for example, fry quality has been reported to show significant G×E interaction effects (Affleck et al. 2008) which was similarly found by Hayes and Thill (2003), who recommended that genotypes should be tested for fry quality over multiple locations. Genetic parameters obtained for fry score in the present study indicated that evaluation in p-rep trials would be appropriate, with broad-sense heritabilities of about 0.8 and correlation coefficients of 0.7 between the C2 and C3 stages (Table 3-5). To test the G×E component (locations and years), further trials are planned for a third year and will be extended to include Pukekohe as well as Lincoln.

Implications for selection

Selection at the initial stages of a potato breeding cycle is often based on the phenotypic appraisal of single plants grown from true potato seed (TPS) or seed tubers, or genotypic evaluation based on progeny families. There is also general agreement amongst potato breeders from previous work that visual selection for a number of economically important traits, such as tuber yield and quality components, is ineffective at the initial seedling and first clonal stage (e.g. Anderson and Howard 1981; Tai and Young 1984; Brown and Caligari 1986; Brown et al. 1987b; Gopal et al. 1992; Love et al. 1997). Traditionally, in a potato breeding scheme, it is typical that testing for quantitative traits such as tuber yield therefore does not begin until at least the second clonal generation when there are enough seed tubers available for formal replicated trials to estimate genotypic effects. At PFR Pukekohe, for example (Fig. 1-3, Chapter 1), seedlings are grown from true potato seed (obtained from selected parental crosses made in the previous year) and planted directly into the field. Visual selections are made on the individual plants (largely for defective tubers and plants), tubers from selected genotypes are collected and these are grown the following year as informal four-tuber clonal plots (C1). These are planted in non-randomised family groups with two check varieties and selection of plots is again based on visual appraisal. This grouping is favoured by breeders because of the ease with which visual within-family comparisons can be made. Alternatively, it is likely that a more formal and improved evaluation step could easily be adopted by choosing to apply a p-rep trial for formal analysis at the C1 stage. This may be an option given the results of the present study, although previous studies may not support this conclusion, e.g. Caligari et al. (1986); Gopal et al. (1992). With this in mind, it should also be noted that the four-plant plots in the present study were C2 (selected from single plants grown from tubers). C1 four-plant plots in the Pukekohe breeders trials are

grown from tubers that have been selected from seedlings. This may incur significant 'maternal' carry-over effects, e.g. Brown (1988), and less reliable results in p-rep trials, which would need to be determined. However, the early stages of a breeding programme expose the greatest amounts of genetic variation and given the evidence for the limited contribution of breeding to the increase in potato yield potential that has been reported over the past century (e.g. Sneep and Hendriksen 1979, pp.144, 421; Douches et al. 1996), such a strategy may be pertinent. This would probably only be feasible, however, if machine harvesting and some degree of automated and rapid phenotyping for yield could be implemented on a small-plot basis. Four is probably the minimum plot size that will be acceptable at C1, and five or six the maximum because of the shortage of seed tubers obtained from single plants. When testing different sized potato test plots and dimensions, Caligari et al. (1985) did not detect a plot size effect, but there may be genotype × density effects due to interplot competition when testing candidates in small plots and therefore bias in the estimate of monoculture performance (Bos and Caligari 2008). Connolly et al. (1993) found that competitive effects in potato trials did not cause a re-ranking of individuals and therefore did not affect selection decisions, but the adjustment for these effects gave an improved prediction of candidate performance in monoculture.

Presently, the main objective at the C1 stage therefore is to reduce the number of genotypes to a manageable size by selecting genotypes on visual preference and eliminating those with obvious faults such as malformed or poorly conformed tubers, or defects such as a propensity to chain tuberise. Visual selections at C1 are then grown at C2 in a replicated trial for formal evaluation at a single location, typically two replicates and 12 plants per replicate. The main recommendation resulting from this study would be to implement a p-rep trial at C2 and to distribute the seed tubers for testing over two locations, namely Pukekohe and Lincoln, which are the two main trial sites in the PFR programme. Candidates are currently tested over the two sites from the C3 or C4 stages onwards. This would allow more genotypes to be tested at the C2 stage without increasing the number of plots required for testing. In previous unpublished studies, genetic correlations for marketable yield between Pukekohe and Lincoln trials were found to range from approximately 0.3 to 0.7. To identify broadly adapted clones, Haynes et al. (2012a) recommended the distribution of tubers to multiple locations in the eastern USA at the early stages of a potato selection programme.

3.6 Conclusions

P-rep trials provide an opportunity to increase the number of genotypes that are tested in a single site and also allow an extension of trials to multiple locations for MET testing at an earlier stage than is currently practised. Based on empirical trials and simulation, results indicate that p-rep trials in the New Zealand potato improvement programme will increase the rate of genetic gain for a number of tuber yield and quality components. Further advantages are possible if material can be distributed across trial sites, i.e. multiple locations, at an earlier stage. It is concluded that full replication at the early stages in a selection programme may be sub-optimal and the use of p-rep designs should be considered as a means to improve the selection efficiency of potato breeding.

4 Genetic variance models for the evaluation of resistance to powdery scab (*Spongospora subterranea* f. sp. *subterranea)* from long-term potato breeding trials

4.1 Summary

Breeding for resistance to soil-borne powdery scab in potato is an important component of the integrated management of this disease. Different genetic variance models within a mixed model framework were applied to data from long-term potato breeding trials, for the genetic evaluation of breeding lines. The multi-environment trial (MET) data came from 12 growing seasons ("years", synonymous with environments) of New Zealand field trials screening for resistance to powdery scab on potato tubers. Pedigree information on a total of 1031 genotypes was available. Additive components of the genetic effects were important with narrow-sense heritability estimates (and 95% confidence intervals) from single-year analyses ranging from 0.26 (0.20, 0.44) to 0.57 (0.53, 0.85). Spatial components estimated from the residual plot effects were not important for most years and even when they were significant, estimates were small. In MET analyses, different variance structures for the genetic effects were tested; a homogeneous correlation model (CORH) gave a better fit to the data than a factor analytic FA*k* model of order (*k*), 1 and 2. The year-to-year genetic correlation estimate from CORH was 0.81 and compared with a range of 0.59 to 0.95 estimated from the FA1 model. There was no strong evidence of non-additive genetic effects with zero or boundary estimates for most years. Models which included the pedigree provided a better fit to the data than models that did not include this relationship information. There was no evidence for genetic improvement in resistance for powdery scab on tubers in the breeding population studied. This suggests that selection pressure for resistance in the past has been weak and greater consideration should be given to selecting parents on empirical breeding values to genetically improve breeding populations for resistance to powdery scab.

4.2 Introduction

Powdery scab is a soil-borne disease of potato tubers (*Solanum tuberosum* L.) caused by the Cercozoan pathogen *Spongospora subterranea* f. sp. *subterranea*. This disease is a problem in New Zealand and Australia, with increasing global prominence in many other potato-producing regions (Merz 2008; Merz and Falloon 2009). The pathogen causes surface lesions

on potato tubers (Fig. 4-1a). Affected tubers can severely reduce quality, consumer acceptability and productivity for all sectors of the industry; seed tubers will not have the required high health status for sale and both the marketable and factory yields of potatoes (grown for human consumption) will be reduced. Zoospores of the pathogen also infect root epidermis cells, and root galls later develop (Fig. 4-1b). Root infection adversely affects plant productivity by reducing plant dry weight and tuber yield (Falloon et al. 2005; Shah et al. 2012). Chemical control using pesticides applied to the seed tuber or soil has partial efficacy but is costly, and soil applications of synthetic pesticides are becoming less acceptable because of environmental concerns (Merz and Falloon 2009). Once *S.subterranea* is established in the soil, long cropping rotations are required for potato to avoid the disease because of the longevity of resting spores (Falloon et al. 2011) . There are therefore both economic and environmental incentives for the potato industry to develop new cultivars that are resistant to powdery scab, which is considered an important part of an integrated disease management strategy (Genet et al. 2004; Falloon 2008).

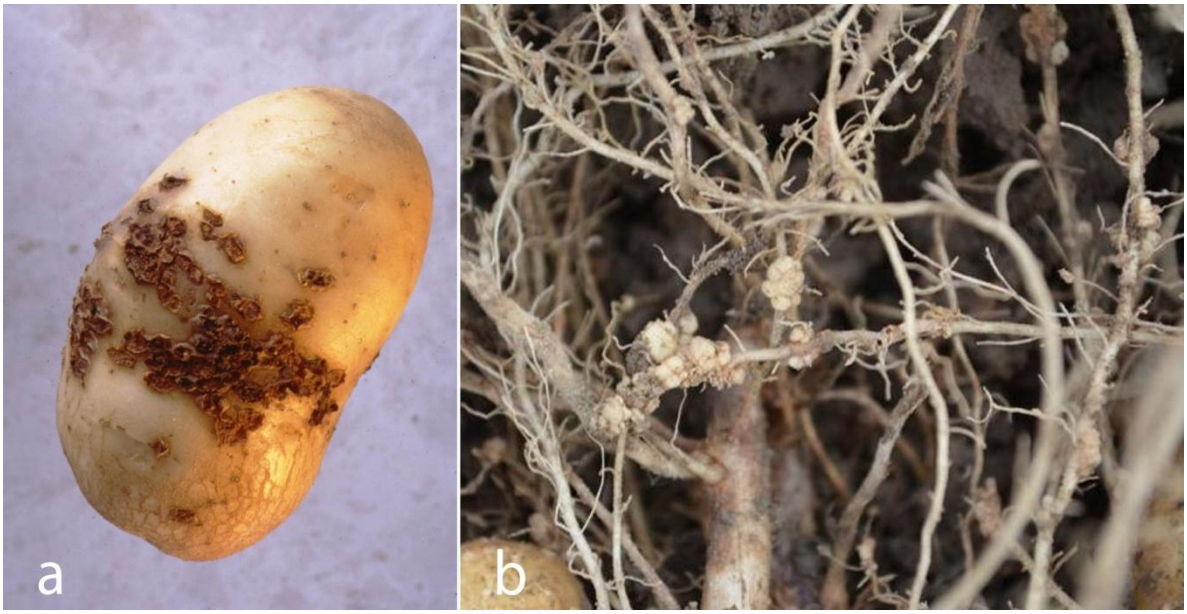

Figure 4–1 (a) Powdery scab lesions on the surface of a potato tuber (photo: Robert Lamberts, PFR) **(b)** *Spongospora* galls on potato roots (photo: Richard Falloon, PFR)

The screening of potato breeding lines and potential cultivars for resistance to powdery scab of tubers has been an important component of the New Zealand potato improvement programme since 1991. Identifying resistant lines is important for the selection of new cultivars for industry and also for the selection of new parents to introduce into the programme, with the intention of increasing the level of resistance in the breeding population by recurrent selection. Studies on the genetic basis for resistance to tuber infection have been

limited but indicate that genetic control is polygenic. In clonal studies on advanced lines and varieties, individual genotypes could be discriminated in terms of their response to tuber infection on a continuous range of susceptible to resistant clones (Bhattacharya et al. 1985; Falloon et al. 2003). Wastie (1991) found that progeny resistance of tuber infection in seedling families was correlated with the phenotypic resistance of parents, indicating the presence of heritable variation and the potential for early-stage selection for resistance in a breeding programme. The genetic relationship between resistance to tuber powdery scab and root infection is also not well understood, with several studies indicating that genetic control for the two may be independent (van de Graaf et al. 2007; Baldwin et al. 2008; Merz et al. 2012). Harrison et al. (1997) suggested that wild *Solanum* species such as *S. x curtilobum* and *S.tuberosum subsp. andigena*, that have shown high levels of resistance to powdery scab, may be useful sources of genetic variation for breeding programmes.

In genetic evaluation, estimates of variance components are required by crop breeders to obtain genetic parameters to plan breeding strategies and to predict breeding values to identify superior parents and breeding lines. Linear mixed models have received increasing attention in recent years for the evaluation of plant breeding multi-environment trials (METs). These models provide a more realistic representation of the underlying random and error components compared with traditional linear models, by accommodating heterogeneous variances and correlated information from trials and/or relatives through a pedigree (Balzarini 2001; Smith et al. 2005; Piepho et al. 2008a). METs are an important part of plant improvement programmes, to assess the response of genotypes to biotic and abiotic pressures that will invariably fluctuate between different growing seasons and locations. Measurement of genotype × environment (G×E) interaction in crop breeding trials aims to identify high-performing genotypes that are either specific to particular environments, or demonstrate a broad stability across a range of environments. A common approach to modelling G×E effects of genotype performance, following Falconer (1952; cited by Falconer and Mackay 1996), is to consider different environments as different traits. The long-term powdery scab screening trials at The New Zealand Institute for Plant & Food Research Limited (PFR), Lincoln, can be considered as a MET in which every growing season is, in effect, a different environment. At present, the evaluation of breeding lines in the early stages of the PFR potato selection programme assumes independence between genotypes tested within each annual trial as well as independence of genotypes tested across trials and growing seasons. This is common practice in many plant breeding programmes, including potato, but is not a particularly efficient use of data. There are various genetic variance models to describe the

structure of the genetic (co)variances within the mixed model framework in MET evaluation (Smith et al. 2001; Crossa et al. 2006; Kelly et al. 2007). One of the simplest, for example, is a homogeneous covariance structure that models different within-trial variances and the same genetic correlation between trials. The most general form (as it attempts to most closely represent the true underlying structure) is an unstructured covariance matrix that contains $t(t + 1)/2$ distinct parameters, where t is the number of trials. This approach is not often feasible under a standard REML-based procedure, even when t is not particularly large, as genotype effects are often highly correlated between some trials resulting in singular variance matrices (Kelly et al. 2009; Meyer 2009). A factor analytic (FA) approach has been considered in plant breeding as a parsimonious alternative to the unstructured form of the genetic variance matrix (Piepho 1998; Smith et al. 2001). FA methods aim to simplify the G×E interaction effects into a small number of unobserved latent variables that attempt to explain most of the interaction. Such models have been shown to be computationally efficient and robust, giving good approximations (Thompson et al. 2003; Kelly et al. 2007; Burgueño et al. 2011), and have been applied, for example, to the MET analyses of cane sugar content in sugarcane (Oakey et al. 2007), yield in wheat (Crossa et al. 2006; Burgueño et al. 2007), yield and oil content in canola (Beeck et al. 2010) and fruit weight in mango (Hardner et al. 2012).

This paper compares various forms of the genetic variance matrix following the general approach of Smith (2001) for the quantitative genetic evaluation of resistance to powdery scab in potato. In addition to comparing variance matrices, the usefulness of including pedigree information was also explored, following the approach of Oakey (2007). Comparisons of model fit with variance matrices were made for data collected from early-stage selection trials for the New Zealand PFR potato breeding programme. Trials were carried out across 12 growing seasons and included a total of 1031 tested genotypes. Including the pedigree information allowed retrospective evaluation of parents based on their empirical breeding values (EBVs), i.e. their ability to transfer powdery scab resistance to progeny, and an assessment of the genetic improvement of resistance to powdery scab in the PFR breeding population.

4.3 Materials and Methods

Data

Data were collected from 12 growing seasons of field trials planted each September from 1998 to 2010 (henceforth designated as "years"). No data were available from 2006. The plant material consisted of 1031 breeding lines (henceforth used interchangeably with "genotypes") originating from crosses of selected parents (with the absence of any formal mating design). The genotypes had already been through one or two clonal generations of selection for tuber yield and quality traits. Many genotypes were related to others planted in the same year or across different years, for example, as full or half siblings. Nine hundred and ninety-nine out of 1031 tested genotypes had both parents recorded in the pedigree. There were 184 female parents and 80 male parents, 48 of which had acted as both male and female parents. Of the tested genotypes, 901 and 948 of both maternal and paternal grandparents respectively were recorded. Forty-eight tested genotypes had been used as parents and eight were grandparents.

Table 4–1 Concurrence of genotypes across 12 years of powdery scab field trials; diagonal entries are the number of genotypes tested in individual years

	1998	1999	2000	2001	2002	2003	2004	2005	2007	2008	2009	2010
1998	160	69	39	4	2	8	5	2	2	2	2	2
1999		140	64	6	2	6	2	2	2	2	2	2
2000			131	39	2	6	2	2	2	2	2	2
2001				140	37	16	8	3	2	2	2	2
2002					122	43	25	14	3	2	2	2
2003						122	45	26	6	2	2	2
2004							113	52	11	2	2	2
2005								113	17	6	4	2
2007									93	32	16	2
2008										136	60	42
2009											136	42
2010												146

There was reasonable concurrence between adjacent years of genotype entries for most pairs of years (Table 4-1). Only two genotypes, namely the New Zealand bred cultivars 'Iwa' and 'Gladiator', were represented in every year. Improving connectedness between pairs of years will enhance the reliability of estimates of year to year genetic correlations and the

accuracy of EBVs. Genetic links other than those between adjacent years were likely to be improved through pedigree relationships.

The trials were grown at The New Zealand Institute for Plant & Food Research Limited (PFR) farm, Lincoln, Canterbury, New Zealand (latitude 43^{o} 38` S, longitude 172$^{o}$ 29` E). The soil at the site was a Paparoa silt loam. Trials were planted on the same field site for three consecutive years. After 3 years, trials were then moved to an adjacent site in the same field. Crop management regimes were consistent for all years. After soil preparation, each trial was planted as a randomised complete block design (RCBD). A single tuber of each genotype was planted in each plot, and each genotype was replicated six times. *Spongospora subterranea* inoculum was made up of macerated potato tubers with severe powdery scab. Each tuber plot was inoculated with approximately 100 ml of inoculum and tubers were then covered with soil. Planting was in mid-September each spring. Moist soil conditions were maintained by regular irrigation through the growing season, particularly during the tuberisation and post-tuberisation periods, to encourage powdery scab development. Plots were harvested in early March after the natural senescence of foliage. All tubers greater than 30 g from each plot were harvested and washed free of soil. Tubers were then assessed for powdery scab severity with a single score assigned to each plot based upon visual assessment of all tubers. Tuber assessment was scored on an ordinal 0 to 9 scale, where 0 = no visible lesions and 9 = complete surface area covered by powdery scab lesions (Fig. 10-3, Appendix III). This scale was adapted from the scoring scheme described by Falloon, (1995). All tuber assessments in the 12 trials were made by the same assessor. The pedigree was built on records from historic PFR field books and a publically-available potato pedigree database (van Berloo et al. 2007).

Model selection

To establish an appropriate statistical model, the first stage was to identify the important non-genetic terms following Beeck et al. (2010). Single trials were analysed to establish the importance of terms for each i^{th} trial, estimated with the general form of the linear mixed model:

$$\mathbf{y} = \mathbf{1}m + \mathbf{Z_1 b} + \mathbf{Z_2 g} + \mathbf{e}$$

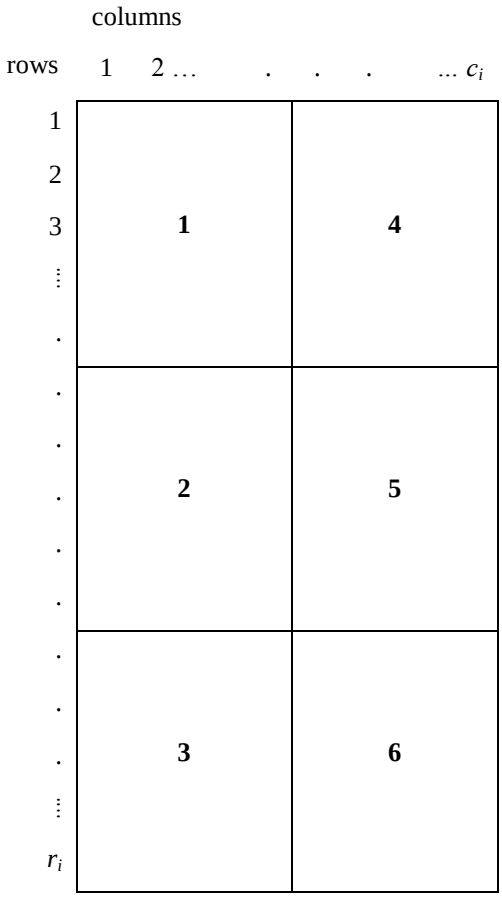

Figure 4–2 Illustrative example of the layout of a randomised complete block design powdery scab assessment trial (*i*) with six replicates in r_i rows and c_i columns

where **y** is the *n x 1* vector of observations for powdery scab severity score, *m* is the overall trial mean as a fixed effect, $\mathbf{b} \sim N(0,\mathbf{I}\sigma_b^2)$ is the *q x 1* vector of random design factors i.e. block or row and/or column effects, $\mathbf{g} \sim N(0,\mathbf{I}\sigma_g^2)$ is the *w x 1* vector of random genetic effects, $\mathbf{Z_1}$(*n x q*) and $\mathbf{Z_2}$ (*n x w*) are known incidence matrices that relate the phenotypic observations to their corresponding vectors, $\mathbf{e} \sim N(0,\mathbf{I}\sigma_e^2)$ is the *n x 1* vector of random error terms and **I** is the appropriate *q x q*, *w x w* or *n x n* identity matrix. The error term included an appropriate spatial model to account for local-scale heterogeneity. Various forms of spatial model are possible. However, a separable autoregressive process of order one (AR1) was the only form considered in this study, as this has been shown to provide an adequate variance structure for local spatial trend in crop breeding trials (Smith et al. 2001), and follows the approach described by Gilmour et al. (1997). Each trial (*i*) comprised a rectangular array of r_i

rows by c_i columns of n_i plots ($n_i=r_ic_i$), as illustrated in Figure 4-2. The best fitting model was selected as the preferred model on the basis of the Akaike information criterion (AIC) goodness-of-fit test: AIC = -2(*logl* -*p*) where *logl* is the REML estimate of the log-likelihood and *p* is a penalty term representing the number of variance parameters fitted.

The model was then developed to incorporate the pedigree, with single trials re-analysed so that the vector of random genetic effects, $\mathbf{g} \sim N(0,\mathbf{I}\sigma_g^2)$, was replaced by the vector of random additive genetic effects, $\mathbf{a} \sim N(0,\mathbf{A}\sigma_a^2)$, where $\mathbf{A}\sigma_a^2$ is the variance-covariance matrix of the additive genetic effects (breeding values), with **A** as the numerator relationship matrix. Variance components, including the additive variance, and a narrow-sense (h^2) heritability for powdery scab resistance were estimated for each trial. In general, heritability in the narrow-sense was obtained from:

$h^2 = V_A/V_P = \sigma_a^2/(\sigma_a^2 + \sigma_b^2 + \sigma_e^2)$; where V_A is the additive genetic variance, V_P is the total (phenotype) variance, σ^2_a is the variance of the additive variety effects, σ^2_b represents the variance of the appropriate design factor(s) (e.g. block or row and/or column effects), and σ^2_e is the random error. Confidence intervals (95%) for narrow-sense heritabilities were estimated by jackknifing so that each sample was generated with the n_j genotype removed. The sampled data was then reanalysed to provide the j^{th} partial estimate and the j^{th} pseudovalue of the heritability, and the vector of pseudovalues was used to approximate the confidence interval.

The single trial model was then extended to a multivariate MET analysis by stacking the vectors for the 12 years $\mathbf{(y_1,y_2,\ldots,y_{12})}$ of phenotypic observations:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_{12}\end{bmatrix} = \begin{bmatrix}\mathbf{1}_1 m_1\\ \mathbf{1}_2 m_2\\ \vdots\\ \mathbf{1}_{12} m_{12}\end{bmatrix} + \begin{bmatrix}\mathbf{Z_{1_1}} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{Z_{1_2}} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z_{1_{12}}}\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\\ \vdots\\ \mathbf{b}_{12}\end{bmatrix} + \begin{bmatrix}\mathbf{Z_{2_1}} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \mathbf{Z_{2_2}} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z_{2_{12}}}\end{bmatrix}\begin{bmatrix}\mathbf{g}_1\\ \mathbf{g}_2\\ \vdots\\ \mathbf{g}_{12}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\\ \vdots\\ \mathbf{e}_{12}\end{bmatrix}$$

and random effects were assumed to follow a multivariate normal distribution with means and variances defined by:

$$\begin{pmatrix}\mathbf{b}\\ \mathbf{g}\\ \mathbf{e}\end{pmatrix} \sim N\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{bmatrix}\mathbf{B_0}\otimes\mathbf{I_b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G_0}\otimes\mathbf{I_g} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{R_0}\otimes\mathbf{I_n}\end{bmatrix}\right]$$

where **0** are null matrices. $\mathbf{B_0}$, $\mathbf{G_0}$ and $\mathbf{R_0}$ are covariance matrices for design factors (block, row, column), genetic and residual effects, respectively, and $\otimes$ is the direct (Kronecker) product. The matrix $\mathbf{B_0}$ is a diagonal matrix of scaled identity matrices and plot error effects $\mathbf{R_0}$ are assumed to be block diagonal. The same site was used for trials in three consecutive years, but the independence of design factors (block, row, column) and plot errors, so that $\mathbf{B_0}$ = diag($\mathbf{B_{0i}}$) and $\mathbf{R}$ = diag($\mathbf{R_i}$) was assumed. This seemed reasonable as trial alignment was likely to change from year to year so that actual block, row, column and plot positions were relocated. Local spatial trend where necessary, as identified in the single trial analyses outlined previously, was specified through $\mathbf{R}$ as a separable autoregressive (AR1) process (Gilmour et al. 1997) and following Smith (2001), with rows within columns:

$\mathbf{R_i} = \sigma_i^2 \boldsymbol{\Sigma c_i} \otimes \boldsymbol{\Sigma r_i}$

where σ_i^2 is a scale parameter and $\boldsymbol{\Sigma c_i}$ and $\boldsymbol{\Sigma r_i}$ are the $c_i \times c_i$ and r_i x r_i correlation matrices, respectively, for the column and row dimensions of the trial, *i*.

The independent genetic component, $\mathbf{G_g} = \mathbf{G_{0g}} \otimes \mathbf{I}$, was then partitioned into additive and non-additive components to test the importance of additive and non-additive genetic effects, as outlined by Oakey et al (2007) and Kelly et al. (2009). The assumption was that the variance matrix for the additive genotype effects was a two-way table of genotype by environment effects which had the separable form $\mathbf{G_a} = \mathbf{G_{0a}} \otimes \mathbf{A}$; where $\mathbf{G_{0a}}$ is the symmetric and positive definite matrix of additive variances and covariances between environments and **A,** the numerator relationship matrix, is the symmetric and positive definite (co)variance matrix between genotypes. Similarly, non-additive effects were considered as a two-way table of genotype by environment effects with the variance structure assumed to have the separable form $\mathbf{G_n} = \mathbf{G_{0n}} \otimes \mathbf{I}$ with independence between non-additive genetic components. This provided the most general form of the models fitted, where the genetic component of the model was partitioned into:

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{n} \end{pmatrix} \sim N\left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{bmatrix} \mathbf{G_{0a}} \otimes \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{G_{0n}} \otimes \mathbf{I} \end{bmatrix}\right]$$

Using the important non-genetic terms identified within each trial, various forms of the genetic variance matrix were then tested against each other. $\mathbf{G_0}$ is the genetic variance matrix with the diagonal elements representing genetic variances for each trial and the off-diagonal elements representing genetic covariances between pairs of trials (synonymous with environments or years). Definitions of the forms of $\mathbf{G_0}$ were as follows:

SIMPLE: all variances within trials were assumed to be equal and all pairwise covariances between trials were assumed to be independent and therefore zero; **DIAG**: variances within trials were assumed to be different and all pairwise covariances between trials were assumed to be zero; **CORH**: variances within trials were assumed to be different and a constant non-zero correlation was assumed between all pairwise combinations of trials; **FA*k***: factor analytic, (Piepho 1997; Smith et al. 2001), with common factors identified (as the leading principal components) and residuals, or specific (lack-of-fit) variances, so that the genetic (co)variance matrix is, $\mathbf{G_0} = \mathbf{\Lambda\Lambda'} + \mathbf{\Psi}$, with $\mathbf{\Lambda}$ a (t x k) matrix of common factors (or environmental loadings) and $\mathbf{\Psi}$ a (t x t) diagonal matrix of specific variances; **RR*k***: (fully) reduced rank analogous to the FA*k* model but all specific variances were assumed absent and fixed at zero (Meyer 2009); **US**: unstructured G_0 with different variances within trials and different covariances between all pairwise combinations of trials.

All models were again tested with and without the pedigree. The analyses of the data were undertaken using ASReml (Gilmour et al. 2006) and R (R Development Core Team 2012), with the mixed models fitted using ASReml-R (Butler et al. 2009). AIC was used as the test criterion for selection of the best model of the various forms of G_0. These were also compared by simulating the response to selection (Piepho and Möhring 2007), based on 1000 simulation runs. Breeding values estimated from the data were assumed to be the true breeding values. For each simulation run, residuals were resampled (with replacement) and added to the fitted values. These new data were then reanalysed to give the best linear unbiased predictors (BLUPs) of genotype effects. At selection fraction *s*, the top *s*100% genotypes, based on the BLUPs, were identified. A response to selection was then computed as the difference between the mean of the true breeding values of these top genotypes and the mean of the breeding population.

Prediction of breeding values

For each model, empirical breeding values (EBVs) were obtained for all genotypes in the pedigree from the best linear unbiased predictors (BLUPs) of breeding line effects, as, for example, outlined by Smith et al. (2005, 457–459), with all years having equal weighting. For the evaluation of the early-stage powdery scab trials, the aim is to rank genotypes for selection. It was therefore appropriate for breeding lines to be considered as random effects. This contrasts with the comparison of genotypes in late-stage trials, when differences between pairs of varieties are of greater interest. In this case, it is considered that variety effects may be regarded as fixed and genotype values obtained from best linear unbiased

estimators (BLUEs) of variety effects, because the BLUP of a specific difference is biased (Smith et al. 2005). Variance parameters are unknown and replaced in the mixed model equations with those estimated from the data, giving empirical BLUEs and empirical BLUPs (respectively, E-BLUEs and E-BLUPs). To detect a genetic trend in powdery scab resistance across years, a mean EBV of tested breeding lines was obtained for each year from the best-fitting variance model, where the cohort for each year was made up of breeding lines in their first year of test. Parametric bootstrapping was used to obtain estimates of the 95% confidence intervals of EBVs for each year.

4.4 Results

Data

The phenotypic means of powdery scab severity scores were low for all years (Table 4-2). There was a poor expression of phenotype in 2002, with severity scores ranging from 0 to 5 resulting in a reduced variance. Non-normal distribution of phenotypic observations for powdery scab severity was found in some years, namely 2002, 2009 and 2010 (Fig. 4-3). All distributions tended to be positively skewed, particularly in 2010 when the mean was 1.71 and the median was 0.

Table 4–2 Summary of powdery scab severity scores from field trials, 1998 to 2010

Year	Observations	Mean	Median	Standard deviation	Range	Dimension (row × column)
1998	972	2.13	2.0	1.65	0,8	108 × 9
1999	864	3.24	3.0	1.80	0,8	96 × 9
2000	810	2.71	2.0	1.77	0,7	90 × 9
2001	864	2.62	2.0	1.59	0,8	96 × 9
2002	756	1.45	1.0	1.10	0,5	84 × 9
2003	756	2.93	3.0	1.48	0,7	84 × 9
2004	702	3.03	3.0	2.08	0,8	78 × 9
2005	702	3.97	4.0	2.19	0,9	39 × 18
2006	-	-	-	-	-	-
2007	630	2.90	2.0	2.02	0,8	126 × 5
2008	840	2.39	2.0	2.03	0,8	168 × 5
2009	840	2.16	2.0	2.07	0,9	168 × 5
2010	900	1.71	0.0	2.29	0,9	90 × 10

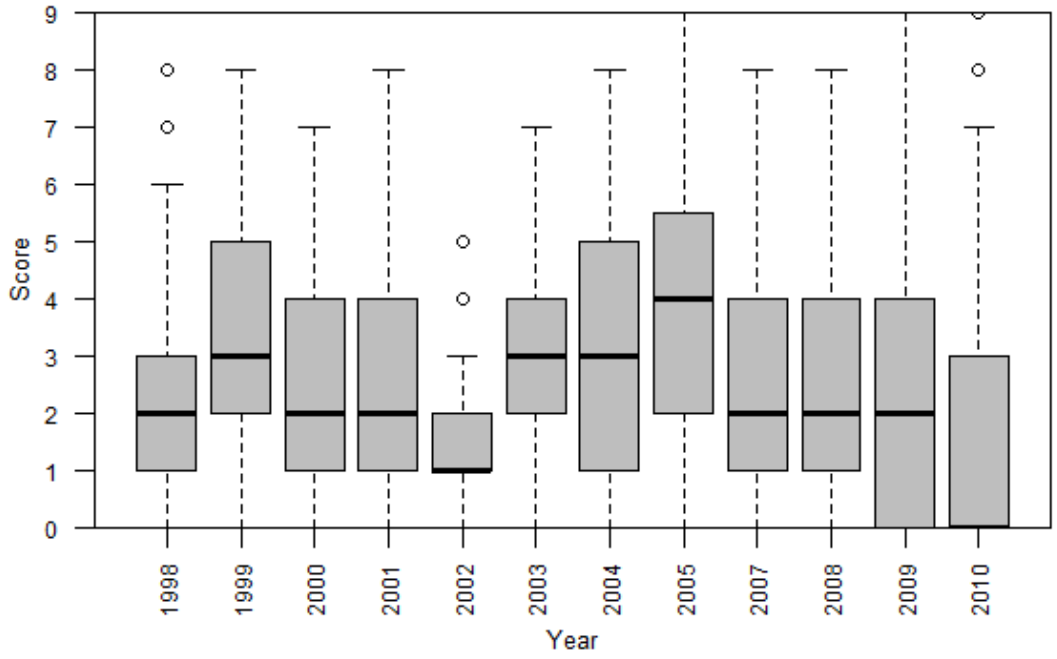

Figure 4–3 Box plots of observed powdery scab severity scores from field trials, 1998 to 2010

Single trial analyses

From the visual interpretation of quantile-quantile plots from single trial analyses (Fig. 10-4, Appendix III), assumptions for the normal distribution of residuals were maintained, despite zero score inflation in some years. It was therefore considered appropriate to analyse these data without recourse to an alternative generalized linear mixed modeling (GLMM) approach. Block terms were important in ten years of the trials with the addition of rows and columns only improving model fit in three of these years (Table 4-3). Spatial terms were positive but small, and only required in four of the twelve years with a maximum correlation of 0.13 for rows in 1998 and 0.09 for columns in 2000. Narrow-sense heritabilities were considered to be moderate, ranging from (95% confidence intervals) 0.26 (0.20, 0.44) in 2002 to 0.57 (0.53, 0.85) in 2007. Additive variance was only 0.33 in 2002 (compared with a range of 0.82 to 2.50 for other years). Including the pedigree improved the fit in all years with the exception of 2002.

Table 4–3 Variance and spatial components, and heritability estimates from single trial analyses (including pedigrees) for powdery scab field trials, 1998 to 2010

	Variance components					Spatial components		
trial	additive	block	row	column	residual	AR1(row)	AR1(column)	h^2[95% CI]†
1998[a]	1.40	0.07			1.35	0.13		0.50 [0.43, 0.56]
1999[a]	1.44	0.06	0.01	0.10	1.62			0.45 [0.36, 0.55]
2000[b]	1.37	0.05			1.55		0.09	0.46 [0.43, 0.68]
2001[b]	1.51	0.05			1.20	0.10	0.05	0.55 [0.49, 0.64]
2002[b]	0.33	0.01			0.95			0.26 [0.20, 0.44]
2003[c]	0.82	0.17			1.41	0.11		0.34 [0.26, 0.45]
2004[c]	2.33	0.14			1.86			0.54 [0.50, 0.74]
2005[c]	1.69			0.28	2.89			0.35 [0.28, 0.52]
2007[d]	2.50	0.01			1.90			0.57 [0.53, 0.85]
2008[d]	2.47	0.09	0.17	0.08	1.84			0.53 [0.44, 0.65]
2009[e]	1.67	0.06	0.06	0.15	2.60			0.37 [0.28, 0.52]
2010[e]	2.40		0.21	0.08	2.93			0.43 [0.38, 0.57]

†95% Jackknife confidence intervals; [a,b,c,d,e]Indicates that the same trial area of land was used for disease screening in the years that share the same letter

MET analyses

The preferred non-genetic models for each trial were then applied to a MET analysis by combining all of the data. Using the DIAG form for the genetic variance matrix, $\mathbf{G_0}$, non-additive genetic effects were not significant, with their estimates often zero or constrained at the boundary and therefore dropped from further analyses. Results from the testing of various forms of the $\mathbf{G_0}$ are shown in Table 4-4, with and without the additive genetic relationship matrix **A**. Results could not be obtained from a US genetic variance matrix because of convergence problems. With initial starting values provided, the maximum number of MET trials that could be analysed together with a US matrix was three before running into difficulties.

Table 4–4 Summary of variance models, number of variance parameters and goodness of fit for powdery scab severity scores from field trials, 1998 to 2010

		No. variance parameters		$^{\dagger}\mathbf{G_g}$		$^{\ddagger}\mathbf{G_a}$	
Model	G_0 structure	G_0	Total	§AIC	-2Log-L	§AIC	-2Log-L
1	SIMPLE	1	40	598	17523	409	17335
2	DIAG	12	51	506	17410	335	17239
3	CORH	13	52	162	17064	0	16902
4	FA1	24	63	162	17042	6	16887
5	FA2	35	74	165	17023	10	16868
6	RR1	12	51	257	17161	114	17018
7	RR2	23	62	281	17163	37	16919

†Genetic variance matrix for independent effects. ‡Genetic variance matrix for additive effects. §Expressed as a difference from the best fitting model (Model 3, with pedigree fitted)

Based on AIC, CORH with pedigree (Model 3) was the best fitting model (AIC=17050) and was very similar to FA1 with pedigree (Model 4) (AIC=17056) respectively. RR2 performed better than RR1, but could not compete with the CORH or FA*k* models. The year-to-year genetic correlation estimate for CORH was 0.81 while the year-to-year genetic correlations for FA1 ranged from 0.59 to 0.95 and are illustrated in Figure 4-4. These were compared with correlations estimated from a bivariate model for each pair of years. There was good agreement between many pairs, but the bivariate results showed a greater range (0.41–0.98) with less extreme values found in the FA analysis. This suggested that the retention of a small number of multiplicative (interaction) terms in this model resulted in a shrinkage of back-transformed correlation estimates.

There was no particular year-to-year trend in genetic correlation estimates from the FA1 model (Fig. 4-4), although the correlations between 2002 and 2003, and all other years, were mostly below 0.75. This may have been a consequence of the poorer expression of phenotype during these 2 years, particularly in 2002. All other year-to-year correlations (ignoring 2002 and 2003) were greater than 0.75.

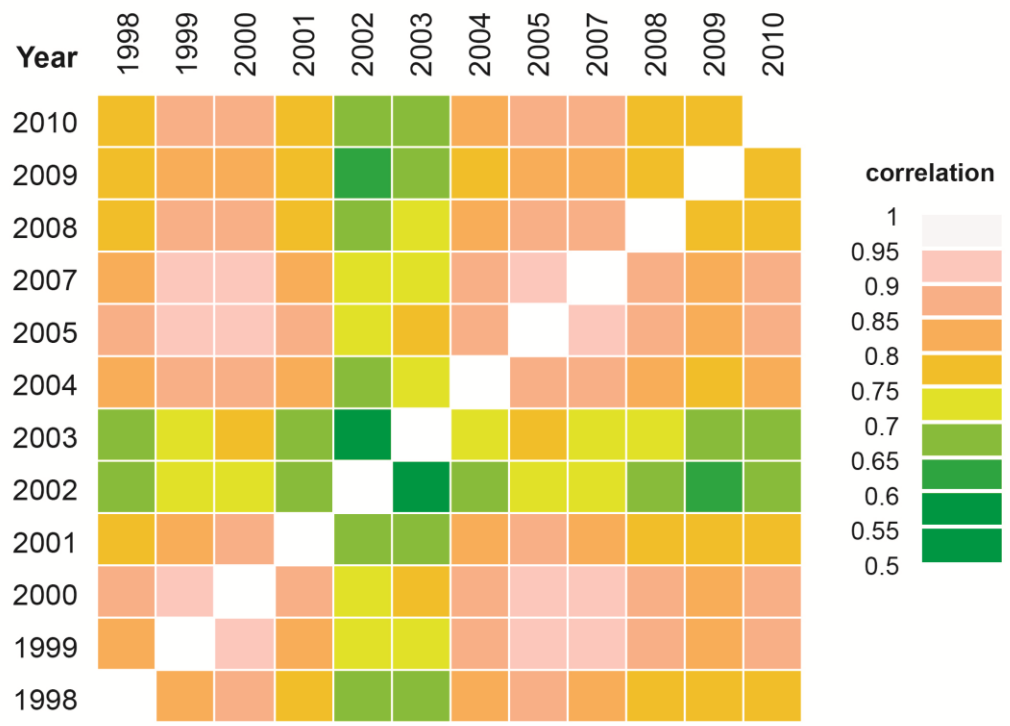

Figure 4–4 Graphical illustration of genetic correlations from FA1 model (Model 4, with pedigree) for powdery scab severity trials, 1998 to 2010

Table 4–5 Responses to selection for DIAG, CORH and FA1 variance models computed by simulation at selection fraction *s*

s	Response to selection (95% confidence interval)		
	DIAG	**CORH**	**FA1**
0.10	0.623 (0.614 0.629)	1.647 (1.601 1.681)	1.628 (1.592 1.661)
0.15	0.558 (0.552 0.563)	1.444 (1.412 1.471)	1.430 (1.400 1.455)
0.25	0.467 (0.462 0.470)	1.151 (1.128 1.173)	1.142 (1.120 1.164)
0.35	0.394 (0.391 0.397)	0.918 (0.900 0.934)	0.912 (0.894 0.927)
0.45	0.336 (0.334 0.338)	0.730 (0.715 0.743)	0.725 (0.711 0.737)
0.55	0.283 (0.281 0.285)	0.571 (0.559 0.581)	0.567 (0.555 0.577)
0.65	0.230 (0.229 0.231)	0.432 (0.422 0.441)	0.429 (0.420 0.438)
0.75	0.174 (0.172 0.175)	0.307 (0.300 0.314)	0.305 (0.298 0.312)

The comparison of BLUPs in Figure 4-5 shows the shrinkage of empirical breeding values estimated from a DIAG genetic variance structure compared with those estimated from CORH. The product-moment coefficient of correlation between CORH and DIAG was 0.73 and between CORH and FA1 was 0.99. From simulation, the response to selection (at different levels of *s*, the proportion of the population selected) from CORH was slightly greater than for FA1 (Table 4-5), although the estimates for FA1 were all within the limits of the 95% confidence intervals for CORH.

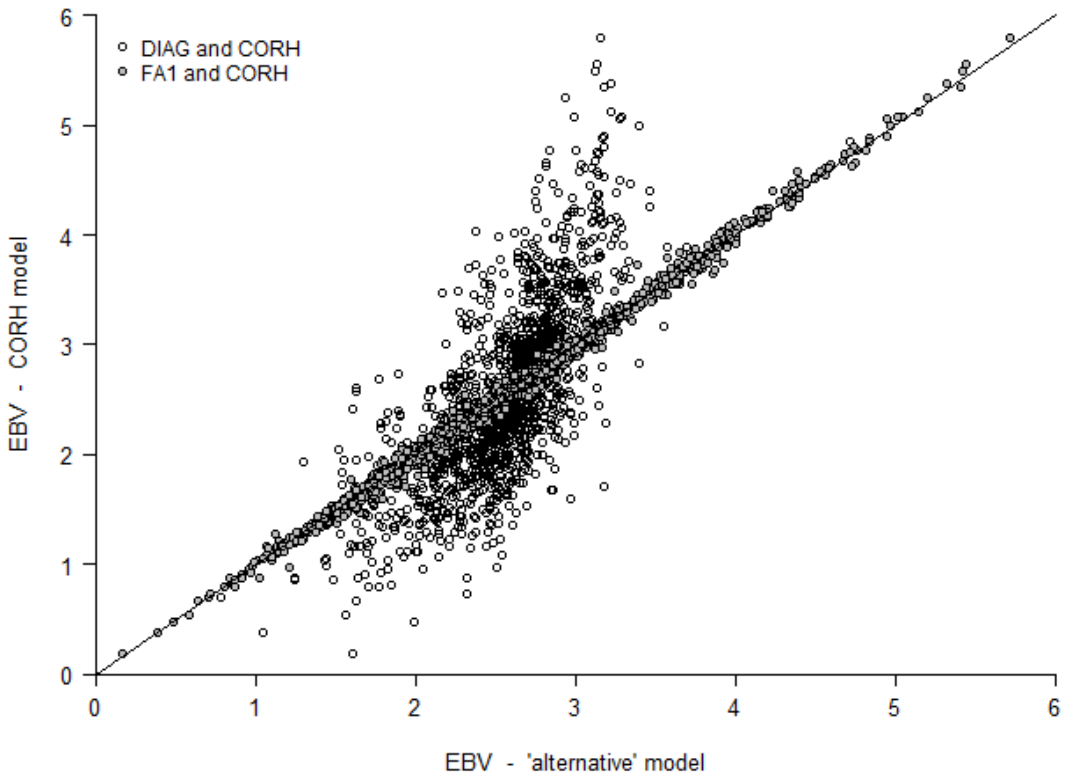

Figure 4–5 Scatterplot of estimated breeding values (EBV), with the 'alternative' model representing DIAG (Model 2) or FA1 (Model 4) against the best fitting model CORH (Model 3), so that: open circles; DIAG and CORH, and filled (grey) circles; FA1 and CORH

Table 4-6 is a (non-exhaustive) list of breeding lines, advanced clones and cultivars of interest that have been used in the past 12 years as parents in the PFR breeding programme. Many of the named cultivars, such as 'Desiree', 'Atlantic', 'Kennebec' and 'Shepody' are popular production cultivars in a number of potato-producing regions of the world. 'Gladiator', a New Zealand-bred cultivar, is highly resistant to powdery scab on tubers, as demonstrated both in New Zealand (Falloon et al. 2003) and Europe (Falloon et al. 2003; Merz et al. 2012). The top ranking genotype is VTN62-33-3, a pollen parent of 'Gladiator', which has been used extensively as a parent in the PFR population in recent years. 'Agria' and 'Kennebec' are popular cultivars widely reported to be susceptible to tuber infection (Harrison et al. 1997). The current study has shown that these cultivars are more likely to transmit disease susceptibility to progeny given their high New Zealand EBVs.

There was no evidence of a genetic improvement for resistance to powdery scab across years in the population of breeding lines tested between 1998 and 2010, as illustrated in Figure 4-6. There was a trend towards greater susceptibility in the population from 2003, but the mean EBVs of the population recovered to 1998 levels in 2009 and 2010.

Table 4–6 Estimated breeding values (EBVs) for powdery scab score; breeding lines, advanced clones and popular named cultivars used as parents in the New Zealand Plant & Food Research potato breeding programme comparing CORH (Model 3) and DIAG (Model 2), ranked on EBVs from Model 3 (resistant to susceptible)

	Parent		CORH (Model3)		DIAG (Model 2)		
Individual	Maternal	Paternal	EBV	SE_{EBV}	EBV	SE_{EBV}	rank
Vtn62-33-3	(V24/20×U.Knight)×Profijt	Vrn I-3 × 'Profijt'	0.37	0.37	1.05	0.27	2
'Gladiator'	B5281-1	Vtn62-33-3	0.70	0.10	0.78	0.09	1
'Fianna'	Konst62-660	Am64-2	1.02	0.33	1.73	0.25	4
761/1	*Unknown*	X61	1.14	0.45	2.34	0.36	14
981/4	Crebella	V394	1.17	0.30	1.62	0.27	3
V394	D47/11	D42/8	1.54	0.18	1.74	0.16	5
'Moonlight'	1463.1	V394	1.70	0.35	1.98	0.24	7
'Nadine'	('Desiree'×'M.Piper')×vrn-seedling	'Pentland Dell'×vrn-seedling	1.72	0.60	2.36	0.32	15
'Van Gogh'	Zpc69-C-239	'Gloria'	1.79	0.56	2.49	0.35	21
'Lone Ranger'	'Ranger Russet'	V394	1.80	0.62	2.14	0.28	8
2765.5	1463.1	Stage2Blue	1.81	0.33	2.40	0.29	17
'Valor'	'Cara'	93/2/10	1.81	0.43	2.38	0.30	16
3011.6	'Fianna'	Nd860-2	2.03	0.30	2.29	0.28	11
'Glenna'	10223-7	10300-13	2.09	0.63	2.44	0.34	20
813/28	354/7	'Fianna'	2.12	0.37	2.60	0.27	25
'Desiree'	'Urgenta'	'Depesche'	2.17	0.55	2.31	0.32	12
1194/7	282/9	Vtn62-33-3	2.24	0.40	1.78	0.30	6
'Red Rascal'	'Tekau'	'Desiree'	2.30	0.48	2.42	0.29	19
'Tutekuri'	*Unknown*	*Unknown*	2.30	1.04	2.59	0.38	23
759/3	178/4	B113-6	2.31	0.44	2.42	0.33	18
'Tekau'	1584C-10	302.01	2.32	0.67	2.64	0.33	27
940/5	'Rua'	L115-1	2.35	0.24	2.81	0.21	37
'Ranger Russet'	'Butte'	A6595-3	2.39	0.34	2.66	0.26	29
'Laura'	'Rosella'	L6140/2	2.40	0.78	2.62	0.36	26
'Summer Delight'	1858.21	V394	2.43	0.32	2.27	0.26	9
'Markies'	'Fianna'	'Agria'	2.45	0.64	2.28	0.31	10
'Draga'	Svp50-2017	Mpi19268	2.58	0.71	2.55	0.36	22
'Maris Piper'	Y22/6	'Arran Cairn'×'Herald'	2.62	0.82	2.34	0.34	13
'Pentland Ivory'	'Pentland Crown'	'Pentland Dell'	2.62	0.97	2.75	0.37	33
'Maris Bard'	Y15/139	'Ulster Prince'	2.72	0.93	2.68	0.36	30
'Atlantic'	'Wauseon'	'Lenape'	2.73	0.49	2.83	0.34	38
3097.5	'White Delight'	V99	2.85	0.42	2.65	0.32	28
'Fraser'	676.34	'Whitu'	2.89	0.52	2.78	0.33	35
'Purple Heart'	1463.1	Stage2Blue	2.97	0.79	2.59	0.33	24
'Rua'	'Katahdin'	'Harford'	2.99	0.47	2.97	0.27	43
1025/2	'Karaka'	L115-1	3.03	0.37	3.17	0.26	51
'Kaimai'	'Rua'	V394	3.06	0.32	2.72	0.23	32
'Allura'	'Kaimai'	L115-1	3.11	0.79	3.10	0.29	50
'Agria'	'Quarta'	'Semlo'	3.18	0.38	2.78	0.28	34
L118-2	H614-1	*Unknown*	3.21	0.38	3.06	0.31	48
'Golden Miracle'	'Agria'	'Fraser'	3.24	0.70	2.81	0.33	36
1021/1	'Fianna'	L115-1	3.30	0.21	3.02	0.21	45
'Karaka'	002/9	V394	3.38	0.29	2.87	0.21	40
2886.3	'Agria'	2221.12	3.43	0.30	3.03	0.26	47
'Kennebec'	USDAB127	USDA96-56	3.45	0.71	3.07	0.33	49
'Katahdin'	USDA40568	USDA24642	3.52	0.75	3.03	0.31	46
'Kiwitea'	002/9	D42/8	3.53	0.53	2.71	0.31	31
L115-1	H612-3	*Unknown*	3.55	0.20	3.52	0.17	54
'Shepody'	'Bake-King'	F58050	3.71	0.62	2.93	0.35	41
2955.19	'Pacific'	'Fraser'	3.87	0.37	2.93	0.32	42
2958.10	285/1	2116.2	3.97	0.35	2.87	0.31	39
2850.6	'Agria'	'Rua'	4.23	0.48	2.99	0.30	44
'Coliban'	'Kennebec'	V28-12	4.39	0.51	3.25	0.31	53
1155/3	'Kennebec'	'Karaka'	5.13	0.40	3.23	0.29	52

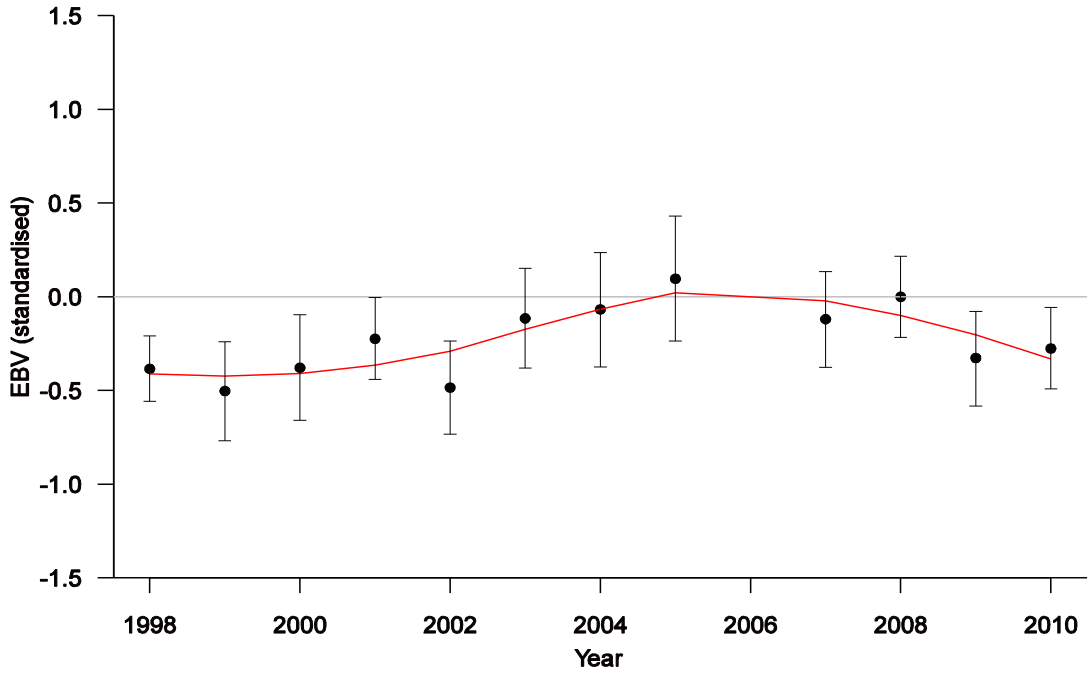

Figure 4–6 Mean (standardised) estimated breeding values (EBV) across years for powdery scab assessment trials, 1998 to 2010, from a CORH model with fitted non-parametric smoothing curve and 95% confidence intervals (from parametric bootstrapping). Presented as the mean EBV of the breeding population as a deviation from the base – the base is the mean EBV of the two common cultivars 'Gladiator' and 'Iwa' and is set to zero. Decreasing values indicate increasing resistance to powdery scab

4.5 Discussion

Many plant breeding programmes have access to historical METs and pedigree data, particularly with the application of versatile plant breeding databases. The genetic evaluation of such data requires robust model selection for accurate estimations of both genetic parameters of traits and breeding values of genotypes. This study confirms the usefulness of models that account for heterogeneity amongst different but related environments in that they are better than simpler versions where this heterogeneity is assumed to be absent.

Factor analytic (FA) models that incorporate pedigree information have been recommended for use in early stage plant breeding METs as they give accurate predictions of treatment effects (Kelly et al. 2007) and are parsimonious alternatives to the unstructured (US) genetic effects model that are computationally tractable (Thompson et al. 2003). The adoption of these approaches by plant breeders within the linear mixed modelling framework however, has been reported to be slow (Smith et al. 2005; Piepho et al. 2008a; Beeck et al.

2010). A number of plant breeding studies have illustrated the modelling of G×E for the analyses of MET data and the benefits of including pedigree information for early stage selection. For instance, studies on canola (Beeck et al. 2010) and barley (Kelly et al. 2009) have demonstrated the superiority of accommodating the pedigree and of FA models over various homogeneous-heterogeneous (co)variance and fully reduced rank (RR) models. The present study also found the FA models were preferred over the RR models (although not as easy to fit), but the application of FA models may not always be necessary for trial data if the heterogeneity of correlated genetic effects between trials is considered to be low. For the long-term trial data analysed in the present study, a more straightforward and parsimonious homogeneous genetic covariance structure (with heterogeneous genetic variances) was found to be adequate when applied to the genetic evaluation of resistance to powdery scab. This was the best fitting model, with a year-to-year genetic correlation estimate of 0.81, which compared with estimates ranging from 0.59 to 0.95 for FA1. From empirical studies under cross-validation, Burgueño et al. (2011) reported that for crop data without G×E crossover interaction, i.e. no re-ranking of genotypes, FA models neither improved nor lost predictive ability (correlation between observed and predicted performance of genotypes) when compared with a simple linear mixed model (homogenous variances, independent covariances between pairs of sites). For other data, when there was crossover interaction, FA models increased the predictive ability by up to 6% compared with the simple model form. FA methods were therefore recommended to model G×E regardless of the type of interaction. Using different variance structures for cross-validation and simulation analyses of maize hybrid MET trials, So and Edwards (2011) found no substantial improvement when including heterogenous genetic (co)variance structures over simpler models. This was attributed to poor genetic links between years. Future work in potato will aim to further augment MET trial and pedigree data with molecular information for the evaluation of traits such as resistance to powdery scab. In wheat MET trials, Burgueño et al. (2012) have shown that G×E models that incorporate both pedigree and marker data improved the predictive accuracy over those that included pedigree information alone.

The magnitudes of narrow-sense heritabilities estimated in the present study were, in general, moderate (i.e. between 0.25 and 0.60) and showed the importance of an additive genetic component in the breeding population. Previous research on the inheritance of powdery scab resistance is limited and does not provide any genetic parameter estimates for comparison. From progeny tests, Wastie (1991) concluded that evaluation of parental phenotype for tuber resistance offered an indication of mean progeny performance, which

suggested that, in the population studied, the additive genetic component was important and the parental phenotype provided a reasonable representation of its breeding value. The narrow-sense heritability estimates published here are based on standard formulae found in quantitative genetics texts for individual trials. Cullis et al. (2006), Oakey et al. (2006) and Piepho and Möhring (2007) have proposed alternative methods to calculate heritability in a MET setting. They argue that heritability estimates from standard methods (although useful in their own right) may not be appropriate to use for the prediction of genetic gain if variance components have been estimated from more complex covariance models, rather than a simple *genotype + residual* model with balanced data. The method of Piepho and Möhring (2007), which directly simulates the response to selection, was used to compare the different variance models tested in the MET analyses. This showed that there was minimal difference between the two models in terms of the response to selection. This supports the result that the CORH model provided only a slight improvement over the FA1 model with regard to AIC as the selection criterion.

The additive relationship matrix **A** used in the individual model in all publically-available software assumes disomic inheritance. This may not be appropriate when dealing with autopolyploid crops, including potato. This is because of double reduction at meiosis whereby sister chromatids can end up in the same gamete. Under simplistic assumptions, such as no past selection, double reduction or inbreeding, the expected additive genetic covariances both of diploid and tetraploid relatives is equivalent (Lynch and Walsh 1998). Despite the acknowledged approximation of the **A** matrix, the inclusion of pedigree information improved model fit for these potato data. This is in agreement with other crops, such as canola (Beeck et al. 2010) and sugarcane (Oakey et al. 2007), which does not meet the general assumptions required of disomic inheritance and obligate (i.e. non-selfing) outcrossing species. Recent work has generalised the relationship matrix and its inverse to better accommodate autopolyploid crop species (Kerr et al. 2012) and should be considered in future work on the genetic evaluation of tetraploid potato.

There was no evidence of any important non-additive effects in the present study indicating that a genetic model with only additive effects is adequate for the evaluation of resistance to powdery scab. Non-additive effects were tested, following Kelly (2009) and Beeck (2010), by simply assuming independence between non-additive genetic components. Oakey et al. (2007) partitioned the non-additive genetic effects into dominance and residual genetic effects. Fitting a non-additive component may remove confounding that could be present, for example, common environment of siblings, and possibly reduce bias in breeding

value estimation. Further, estimates of non-additive effects for individuals could be used to select parental combinations to exploit the non-additive genetic variance (Mrode 2005, p. 193; Oakey et al. 2007). Kelly et al (2009) noted that the estimation of non-additive effects in complex models are likely to often represent only a small proportion of the total genetic variance (e.g. Hill et al. 2008) and are likely to be characterised by large sampling errors.

It is widely reported that there is much variation in the year-to-year incidence and severity of powdery scab and the attributing factors are largely considered to be variations in rainfall and temperature that influence soil conditions and the rate of crop development (reviewed by Harrison et al. 1997; Merz 2008; Merz and Falloon 2009). Screening procedures involving artificial inoculation with pathogens such as *S. subterranea* are expected to inflate the within-trial heritability compared with trials that rely on natural infection. Further, greater stability, or homogeneity, in the estimates of the between-year genetic correlations is likely, particularly if there is some degree of control over the environmental conditions that predispose disease development, such as soil water availability. This may also be the case for many other quantitatively-inherited diseases in various crop species in which progenies are assayed in glasshouse or in field assessments.

As well as the problems associated with tuber infection, such as poor tuber quality and product rejection, *Spongospora* root galling has been shown to reduce root function and plant growth by impeding water and nutrient uptake (Lister et al. 2004; Falloon et al. 2005; Shah et al. 2012). Further research should therefore aim to determine the genetic relationship between tuber infection and root infection. This information will assist in determining an appropriate selection strategy, as simultaneous improvement of both traits is required in the breeding objective. Several studies on clones have found no evidence of correlation between tuber lesion severity and root galling (van de Graaf et al. 2007; Baldwin et al. 2008; Merz et al. 2012), although a positive correlation between the levels of root infection by *S. subterranea* zoosporangium and tuber scab formation has been reported (Falloon et al. 2003). The common potato processing cultivar ‘Russet Burbank’ is known to display few tuber symptoms but often shows symptoms of severe root galling. Root symptoms may go unnoticed but development of *S. subterranea* sporosori (containing resting spores) in root galls will provide reservoirs of inoculum that have long-term consequences for potato production because of the long-term survival of resting spores (Merz et al. 2012). Information on root infection is more difficult (and probably more costly) to routinely collect in early stage testing in a breeding programme, suggesting that powdery scab on tubers would be the preferred trait of the two in a selection criterion. Because of the resources required for

screening roots and tubers, both may be good candidates for molecular-based selection approaches.

Merz (2012) found no evidence of cultivar × location effect for powdery scab resistance in clones over 4 years in a European study. They concluded that multiple-location testing in a cultivar development programme should not be necessary. Evidence indicated limited *S.subterranea* pathotype variation and suggested that the genetic mechanisms for the pathogen to overcome host resistance were probably less dynamic than those associated with airborne pathogens such as *Phytophthora infestans* (which causes potato late blight). It is unclear if there is a sexual phase in the life-cycle of *S.subterranea* but from genomic analysis of diverse field collections, only two genetically distinct groups had been previously identified (Merz 2008). In a more recent study, a total of 19 haplotypes were identified by Gau et al. (2013). Two of these variants were found to be widely distributed globally, while the other 17 were only found in South America. These results suggest that EBV information from trials in, say, New Zealand may be used to aid selection in other locations with a reasonable level of confidence.

The assumption of homogeneity within block and/or row and column factors in the trials appeared to hold as the analysis of individual trials found no evidence of any important localised spatial effects in most years, using an autoregressive procedure of order one. Even where there was an apparent effect for rows (1998, 2001, 2003) and for columns (2000, 2001) the estimated parameters were small and ranged from 0.10 to 0.13 for rows and 0.05 to 0.09 for columns. Their inclusion (or omission) in future evaluations is likely to make very little difference to final selection decisions. However, the effort required to check spatial effects is minimal compared to the effort and resources invested in a trial, and thus it would be advisable to do so. This study considered the separable autoregressive spatial method only, and chose not to investigate other approaches for spatial adjustment (e.g. see Gleeson 1997; Piepho et al. 2008b) as it was not the main purpose of this study. For powdery scab resistance, studies have indicated that the amount of artificially added inoculum in the soil had no effect on the incidence and severity of visual symptoms of lesions on potato tubers (van de Graaf et al. 2007; Shah et al. 2012). Severe infection has been shown to occur when inoculum levels in soil were considered to be low, with no consistent relationship between soil inoculum level and severity of infection (Shah et al. 2012). This may partly explain the apparent lack of important spatial components. Trial sites for the PFR programme are used for three consecutive years and re-innoculated each year. With no reported dose effects, it is unlikely that there will be changes in the importance of spatial components as inoculum

distribution across the sites possibly become more uniform over time (less spatial heterogeneity). In contrast, there may be an expectation of more localised spatial heterogeneity (or 'patchiness') as soil inoculum levels build during this period. No such patterns or trends were evident from these 12 years of data. Soil water content affects powdery scab development on tubers, particularly if water content is high over the tuberization period (Merz and Falloon 2009). The PFR powdery scab trials were regularly irrigated over the growing season, to maintain uniform conditions that are conducive for disease development on tubers.

There was no evidence of genetic improvement for powdery scab resistance – that is, no incremental and consistent decrease of mean EBV in the breeding population over the 12 years of trials. This suggests that selection pressure for the disease has been weak, possibly partly due to a slow and reluctant replacement of older and susceptible parents and long generation intervals in the breeding programme. Further, we have little or no understanding of the genetic relationships between powdery scab resistance and other traits that may feature more prominently in a breeding programme. Powdery scab resistance is recognized by breeders to be an important component of an economic breeding objective, but other traits, such as yield and tuber quality, are likely to exercise much more weight in an (implicit) selection index. Resistant cultivars have been developed, but this has relied on selection of individual clones that display an acceptable level of phenotypic resistance at an advanced stage of selection. It has not been a direct consequence of a population improvement approach. The cultivar 'Gladiator', for example, was developed in the PFR potato breeding programme and has demonstrated a high level of powdery scab resistance in New Zealand and Europe (Merz et al. 2012). Selection of resistant parents based on estimates of their EBVs should help confer resistance to progeny, and therefore be part of a recurrent selection strategy for a population-based improvement of this economically-important characteristic.

4.6 Conclusions

This study has shown that a homogeneous genetic correlation model (with heterogeneous variances) was simpler and more parsimonious than a factor analytic model determined from METs for the genetic evaluation of resistance to powdery scab in potato tubers. Simpler models should not be overlooked in the evaluation of plant breeding data if they can compete with more complex forms. Unnecessarily increasing the complexity of models should be avoided if plant breeders are to routinely adopt the genetic evaluation of MET data in a linear

mixed model framework for early stage selection. The additive component of variation was important and narrow-sense heritabilities were moderate. The year-to-year genetic correlations were generally high. There was no evidence of non-additive genetic effects, and local-scale spatial heterogeneity was not apparent. Further work is required to determine the genetic relationship between tuber powdery scab and *Spongospora* root infection to help devise a comprehensive breeding strategy for resistance to this pathogen. Exploiting correlated information for the estimation of breeding values in disease screening trials should assist breeders to improve the quantitative resistance to powdery scab of potato tubers in breeding populations. The success of this in a multi-trait selection strategy will, of course, ultimately depend on a number of other factors such as its hitherto unknown genetic relationships with other target traits and the relative importance it is afforded in the selection criterion.

5 Selection for tuber yield in a potato improvement programme: trial heterogeneity and genetic variance models in early stage evaluation

5.1 Summary

Genetic evaluation aims to identify genotypes with high empirical breeding values (EBVs) for selection as parents. In this study, 2157 potato genotypes were evaluated for tuber yield using 8 years (21 trials) of early-stage trial data collected from a potato breeding programme between the years 1999 and 2012. Using an individual plant model, spatial parameters to target greater control of localised spatial heterogeneity within trials were estimated and variance models to account for across-trial genetic heterogeneity were tested. When spatial components improved the fit of the model, the correlations of errors for rows and columns were mostly small and negative for marketable tuber yield (MTY) and total tuber yield (TTY), suggesting the presence of interplot competition for plot yield in some years. When testing different genetic variance models for the analysis of multi-environment trial data, a heterogeneous variance-homogeneous correlation model (CORH) was the most favourable variance structure fitted for TTY and CHI (percent marketable yield). There was very little difference in model fit when comparing a factor analytic structure of order 2 (FA2) with either FA1 and CORH structures for MTY, but simulations of the genetic response to selection indicated a possible introduction of bias and over-fitting of the FA2 model. Simulations also indicated that MET evaluation models that use a simple homogeneous genetic correlation structure and also assume localised within-trial spatial independence may make very little difference to the realised genetic gain of potato yield compared with selection from more complex models. The evaluation of potato tuber yield in the early stages of a breeding programme is discussed.

5.2 Introduction

Technological developments in both genetics and agronomy have underpinned significant yield gains in staple crops for many global regions over the past century (Kang 2002a). In New Zealand, for example, records show that potato tuber yield has increased from an average of 11 tonnes per hectare in 1930, based on a total production area of 9400 hectares, to approximately 46 tonnes per hectare in 2007, based on 10850 hectares (FAOSTAT 2013). In many field crops, a high proportion of reported yield gains (on a unit area basis) over the

last century have been attributed to genetic improvement and cultivar development (e.g. Kang 2002b; Kang 2002a; Evenson and Gollin 2003), whereas the contribution of genetics to yield gains in potato is reported to have been small in comparison (Sneep and Hendriksen 1979, pp. 144 & 421; Douches et al. 1996; Walker et al. 2003).

Marketable tuber yield is an important selection criterion and contributes to maximising the (implicit) economic objective in most potato breeding programmes. A number of studies have shown that plant yield and other tuber components measured in the initial stages of selection (seedling or first clonal stage) were poorly correlated when assessed from replicated plots in later clonal generations (e.g. Anderson and Howard 1981; Brown and Caligari 1986; Caligari et al. 1986; Gopal et al. 1992). This has resulted in recommending the testing of seedling families in the early stages to improve selection efficiency (Bradshaw et al. 2009). It is typical that information on tuber yield under either a phenotypic or genotypic (progeny testing) selection strategy is not available until at least the second clonal generation, when there are enough tubers available for establishing formal replicated trials. Under a traditional phenotypic recurrent selection strategy, the use of a promising parental candidate will often be delayed until the breeder has enough confidence in its individual 'production worth' after further years of extensive field trials. Furthermore, the candidate's 'breeding worth' is not necessarily formally evaluated from the performance of its progeny when this information eventually comes to hand. Evaluation of potato genotypes may therefore fail to take advantage of all available information in multi-environment trial (MET) data, which might improve the precision of breeding value estimation in the early stages of testing; trial analyses often assume the independence of genotypes both within and across trials, but these assumptions are not realistic (Smith et al. 2005). Genetic evaluation for yield using trial data and ancestry information can provide predictions of breeding and genetic values from informal mating designs at the early stages of breeders' trials (Oakey et al. 2006; Kelly et al. 2007; Oakey et al. 2007; Beeck et al. 2010).

Observations that the residuals of neighbouring plots in field trials are often more alike than those of non-neighbours have led to the development of a number of statistical approaches to deal with this spatial dependence (Gleeson 1997; Edmondson 2004). These methods attempt to account for small-scale or localised trial heterogeneity that is not accounted for by standard blocking procedures, such as classical randomised complete block or more advanced incomplete block designs, such as row-column arrangements (Basford et al. 1996). An interest in augmented and partially replicated trial designs for early-stage potato

selection in The New Zealand Institute for Plant & Food Research Limited (PFR) potato breeding programme has motivated the investigation of localised spatial effects.

Inclusion of pedigree information and selection based on empirical breeding values is routine in livestock evaluations (Mrode 2005) and in this context is commonly known as the 'individual animal model' (hereafter referred to as an 'individual plant model') when applied in a mixed model framework. Different variance structures can be set up within a mixed model to accommodate the genetic (co)variances that exist among trials or 'environments' in MET data (e.g. Smith et al. 2001; Crossa et al. 2006; Kelly et al. 2007; Meyer 2009), allowing varying degrees of complexity to be modelled. The most general form is an unstructured (US) covariance matrix, which models both heterogeneity of trial variance and different covariances for each pair-wise combination of trials. This approach is often confronted with computational problems because genotype effects are often highly correlated between some trials, resulting in singular variance matrices (Kelly et al. 2009). Alternatively, trial evaluation may be enhanced by fitting a homogeneous covariance structure that models different within-trial variances and the same genetic correlation between trials. This structure may be relatively simple to fit but may not be reasonable when trials are performed over diverse environments and the assumed homogeneous genetic correlation structure does not adequately deal with the genetic heterogeneity that may exist. To overcome the computational difficulties of fitting the US form of the genetic variance matrix, an alternative is the factor analytic (FA) model (Piepho 1998; Smith et al. 2001). To simplify the calculations, the FA approach attempts to confine the genotype-by-environment (G×E) interaction effects into a small number of components (unobserved latent variables) that aim to explain most of the interaction and in this respect is analogous to ordination methods previously developed to study G×E, such as the Additive Main effects and Multiplicative Interaction (AMMI) model (Gauch and Zobel 1988; Crossa et al. 1991). FA mixed models have been applied to early-stage field MET crop trials in, for example, sugarcane (Oakey et al. 2007), cereals (Crossa et al. 2006; Kelly et al. 2009), and canola (Beeck et al. 2010).

The general aim of this study was to identify an appropriate genetic evaluation model for analysing MET potato yield data for early stage selection in a potato breeding programme. The study used 21 trials from 8 years of MET data, which included 2157 genotypes, from the early-stage selection trials of the New Zealand PFR potato breeding programme. Spatial models were investigated for greater control of local-scale heterogeneity within potato field trials and different variance structures were modelled to account for across-trial heterogeneity using four years of MET field data. This enabled the ranking of potential parents according to

their ability to transmit tuber yield to progeny – breeding value information that will be of direct benefit to potato breeding schemes, allowing informed decisions with regard to parent selection for tuber yield.

5.3 Materials and Methods

Data

The potato selection trials analysed for the study were part of the PFR potato breeding programme. Trials were performed over a number of years (1999 to 2003, 2006 to 2007, and 2012) mostly at Pukekohe, South Auckland (37°.12'S 174°.52'E, 141m asl) but with two trials in Palmerston North, Manawatu (40°.20'S 175°.28'E, 30m asl). The 21 trials were designed as early clonal stage two, three and four (C2, C3 and C4 respectively) 'early-main' (EM) crop and 'main' (MN) crop yield trials. Early-main crop trials were planted in mid to late September and harvested in late February, approximately 150 days after planting. Main crop trials were planted in early November and harvested in mid May, with weather conditions sometimes delaying harvest into June. C2 trials were treated as early-main crop trials. Selected genotypes from the C2 stage were entered into main crop (Manawatu: MW) and early-main and main crop (Pukekohe: PK) trials at the C3 and C4 stages. Each trial comprised a rectangular array of rows by columns, typically of 60 to 90 genotypes replicated twice, designed as a Latinized row-column with CycDesigN v4.0 (CycSoftware 2009) and previous versions of the trial design software. Such designs require that each genotype occurs once, at most, in any given row or column across a trial consisting of individual rectangular plots. Rows of a trial array are defined as being parallel to the smaller plot side and orthogonal to the direction of planting. Crop management practice was consistent for all trials. Each plot was made up of 12 tubers in total, planted in a six by two arrangement, with a width of 1.55 m and a length of 2.0 m. Spacing between neighbouring plots on the shorter plot side was 0.58 m and on the longer plot side was 0.77 m. Plot yield was recorded at harvest as both a total tuber yield (**TTY**) and a marketable tuber yield (**MTY**), and converted to t ha^{-1} (metric tonnes per hectare) for analyses. MTY, the trait of most interest in the present study, was the saleable (graded) yield after undersized (less than 80 g) and defective tubers had been removed. Yield was also expressed as the percentage marketable fraction of the total yield and is referred to as the commercial harvest index (**CHI**). These data were logit transformed so that $ln\left(\frac{p}{1-p}\right)$, where p is the proportion of MTY to TTY.

TTY, MTY and CHI were analysed in 21 early-stage potato breeding trials for the estimation of variance components and spatial parameters. Fifteen of these trials (1999 to 2003) showed reasonable concurrence of genotypes across trials and across years. This representative series of early-stage trials were therefore used to test different variance structures to account for trial heterogeneity and to estimate breeding values for potato yield for both (1050) tested genotypes and all genotypes in the pedigree.

Single trial analysis

Single trials were analysed to estimate variance components for each trial, estimated with the general form of the linear mixed model: $\mathbf{y} = \mathbf{1}m + \mathbf{Z_1}\mathbf{b} + \mathbf{Z_2}\mathbf{g} + \mathbf{e}$, where $\mathbf{y}$ is the vector of yield observations, m is the overall trial mean as a fixed effect, $\mathbf{b} \sim N(0,\mathbf{I}\sigma_b^2)$ and $\mathbf{g} \sim N(0,\mathbf{I}\sigma_g^2)$ are vectors of random (non-genetic) design factors, e.g. row and/or column, and genotypic effects respectively, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_e^2)$ is the vector of random error terms. $\mathbf{Z_1}$ and $\mathbf{Z_2}$ are known incidence matrices of the random effects (trial design and genetic effects) and $\mathbf{I}$ are identity matrices.

A randomisation-based approach (the *base* model) was first used in analyses to reflect the experimental design. This included the independent row and column effects and the complete replicate effects if necessary. The *base* model was then compared with an *extended* row × column model that included random row and column effects and row and/or column spatial correlation parameters, as an attempt to better describe local, small-scale heterogeneity. Various forms of spatial model are possible (e.g. Gleeson 1997; Edmondson 2004; Piepho and Williams 2010). A separable autoregressive process of order one (AR1) has previously been shown to provide a suitable variance structure for local spatial trend in annual and perennial crop and tree evaluation trials (Gilmour et al. 1997; Smith et al. 2001; Dutkowski et al. 2002; Gonçalves et al. 2007), increasing precision in the estimates of genotype effects. and, in general, comparing well with alternative linear variance models (Müller et al. 2010). This also follows the general approach described by Cullis and Gleeson (1991) and Gilmour et al. (1997). In spatial analysis, $\mathbf{e}$ can be decomposed into spatially dependent and spatially independent ("nugget") errors. The following matrix showing the pattern of spatially dependent errors, modelled as the AR1 correlation coefficients (ρ) between plots in the same ordered column or row (q) of size n:

$$\boldsymbol{AR1}(\boldsymbol{\rho_q}) = \begin{bmatrix} \mathbf{1} & \boldsymbol{\rho_q} & \boldsymbol{\rho_q^2} & \cdots & \boldsymbol{\rho_q^{n-1}} \\ \boldsymbol{\rho_q} & \mathbf{1} & \boldsymbol{\rho_q} & \cdots & \vdots \\ \boldsymbol{\rho_q^2} & \boldsymbol{\rho_q} & \mathbf{1} & \mathbf{1} & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\rho_q^{n-1}} & \cdots & \cdots & \cdots & \mathbf{1} \end{bmatrix}$$

This is generalised to give the correlation coefficient between plots not located in the same row or in the same column as $\rho^{|\boldsymbol{i-i'}|}\rho^{|\boldsymbol{j-j'}|}$ for plots separated by $|i - i'|$ rows and $|j - j'|$ columns from the direct product **AR1** ⊗ **AR1**. Any large-scale field trend (large-scale dependence or global trend in the mean of errors) present across rows and/or columns was accounted for by fitting fixed linear or polynomial regressions to the spatial coordinates (Gilmour et al. 1997) or fixed linear regressions and cubic smoothing splines (Verbyla et al. 1999). To examine the pattern of variability remaining after fitting a given model, plots of the residuals against row or column (number) conditional on column and row respectively, were examined. Each residual variogram, described as a function of the semi-variance of the difference between the errors of plots arising at a given distance (of rows or columns) apart, was also inspected. From the variograms, global trends were detected by non-stationarity and fixed linear and polynomial global terms were tested using approximate incremental F-tests based on Wald statistics, with non-significant fixed regression terms sequentially dropped from the model. The *extended* model was then revised, with any trend terms added and compared with the *base* model (with the same fixed effects). An *extended* model was therefore considered as a model that included localised spatial error components and/or global field trends, as well as row and column block effects. The best fitting model was selected as the preferred model on the basis of the Akaike information criterion (AIC) goodness-of-fit test: AIC = -2(*logl* - N_p), where *logl* is the REML estimate of the log-likelihood and N_p is a penalty term representing the number of variance parameters fitted. Smaller values of AIC represented a better fitting model. The modelling procedures then incorporated the pedigree. Models were tested using an individual 'plant' model, with $\mathbf{I}\sigma_g^2$ replaced by $\mathbf{A}\sigma_a^2$, the variance-covariance matrix of the additive genetic effects (breeding values), where **A** as the numerator relationship matrix that provides the between-genotype relationship as two times the coefficient of coancestry.

If the *base* model was not the preferred model, then the preferred model (as the *extended* model) was measured against the *base* model for relative efficiency (RE). The RE assessed the improvement in precision using an *extended* model over the *base*, in terms of the average standard error of the difference between genotype means (SED), so that: RE = 100 ×

(SED_{base}/$SED_{extended}$). The changes in ranking using Spearman rank correlations for all tested genotypes and the percentage concurrence of genotypes selected between the *extended* and *base* models (when the top 10 percent of genotypes ranked on EBV were selected from each analysis) were also measured.

Testing variance models for MET analysis

The single trial models were then extended to a multivariate MET analysis by analysing 15 trials $\mathbf{(y_1,y_2,\ldots,y_{15})}$ between 1999 to 2003, (trial PK-EM-00 was excluded, as only a small number of genotypes in this trial were shared with other trials), with data vectors and design matrices constructed as follows:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_{15}\end{bmatrix} = \begin{bmatrix}\mathbf{1}_1 m_2\\ \mathbf{1}_2 m_2\\ \vdots\\ \mathbf{1}_{15} m_{15}\end{bmatrix} + \begin{bmatrix}\boldsymbol{Z}_{1_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \boldsymbol{Z}_{1_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{Z}_{1_{15}}\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\\ \vdots\\ \mathbf{b}_{15}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_{2_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \boldsymbol{Z}_{2_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{Z}_{2_{15}}\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\\ \vdots\\ \mathbf{a}_{15}\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\\ \vdots\\ \mathbf{e}_{15}\end{bmatrix}$$

Random effects were assumed to follow a multivariate normal distribution with means and variances defined by:

$$\begin{pmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{e}\end{pmatrix} \sim N\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{bmatrix}\mathbf{B_0}\otimes\mathbf{I_b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G_0}\otimes\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{R}\otimes\mathbf{I_n}\end{bmatrix}\right]$$

where **0** are null matrices. $\mathbf{B_0}$, $\mathbf{G_0}$ and **R** are covariance matrices for design factors, genetic (additive) and residual effects, respectively, and $\otimes$ is the direct (Kronecker) product. The matrix $\mathbf{B_0}$ is a diagonal matrix of (non-genetic) scaled identity matrices and plot error effects **R** are assumed to be block diagonal. As each trial (t) analysed comprised a rectangular array of r_t rows by c_t columns (n_t=$r_t c_t$), local spatial trend, as described in the single trial analyses outlined previously, was specified through **R** as an AR1 process (Gilmour et al. 1997) and following Smith (2001), with rows within columns: $\mathbf{R_t}= \sigma_t^2[\boldsymbol{\Sigma c_t}\otimes\boldsymbol{\Sigma r_t}] + \sigma_\eta^2\mathbf{I}$

forming the spatially dependent error matrix and independent residual variance, where $\boldsymbol{\sigma_t^2}$ is a scale parameter and Σc_t and Σr_t are the $c_t \times c_t$ and r_t x r_t correlation matrices associated with the coordinates of the column and row layout of the trials, respectively. σ_η^2 is the independent measurement error variance for trial t (the 'nugget' effect) and **I** is an identity matrix.

The assumption was that the variance matrix for the additive genotype effects has the separable form $\mathbf{G_a = G_0 \otimes A}$ (Kelly et al. 2009), where $\mathbf{G_0}$ is the matrix of additive variances and covariances between environments and **A** is the covariance matrix between genotypes – the numerator relationship matrix. Non-pedigree based models were also tested, so that the independent genotype effects were of the form $\mathbf{G_g = G_0 \otimes I}$, where **I** in this particular case is an identity matrix of order *g* (number of genotypes). Using the important non-genetic terms identified from each single trial analysis, four forms of the genetic variance matrix were then compared with each other. $\mathbf{G_0}$ is the genetic variance matrix, with the diagonal elements representing genetic variances for each trial and the off-diagonal elements representing genetic covariances between pairs of trials. Definitions of the forms of G_0 tested are as follows:

SIMPLE: all variances within trials are assumed to be equal and all pairwise covariances between trials are assumed to be independent and therefore zero:

$$\mathbf{G_0} = \begin{bmatrix} \boldsymbol{\sigma^2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma^2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\sigma^2} \end{bmatrix}$$

DIAG: variances within trials are assumed to be different and all pairwise covariances between trials are assumed to be zero:

$$\mathbf{G_0} = \begin{bmatrix} \boldsymbol{\sigma_1^2} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma_2^2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\sigma_t^2} \end{bmatrix}$$

CORH: variances within trials are assumed to be different and a constant non-zero correlation is assumed between all pairwise combinations of trials:

$$\mathbf{G_0} = \begin{bmatrix} \boldsymbol{\sigma_1^2} & \boldsymbol{\rho_{t,t'}} & \cdots & \boldsymbol{\rho_{t,t'}} \\ \boldsymbol{\rho_{t,t'}} & \boldsymbol{\sigma_2^2} & \cdots & \boldsymbol{\rho_{t,t'}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\rho_{t,t'}} & \boldsymbol{\rho_{t,t'}} & \cdots & \boldsymbol{\sigma_{15}^2} \end{bmatrix}$$

where $\boldsymbol{\rho_{t,t'}}$ represents a constant correlation of additive genetic effects between trial *t* and trial *t'*.

FA*k*: factor analytic, a parsimonious approximation to the US genetic (co)variance matrix (Piepho 1997, 1998; Smith et al. 2001), which identifies common factors (as the leading principal components) and residuals, or 'specific variances', and is given by: $\mathbf{G_0 = \Lambda\Lambda' + \Psi}$; where $\mathbf{\Lambda}$ is a (*t x k*) matrix of environmental loadings (or common factors):

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Phi_{11}} & \boldsymbol{\Phi_{12}} & \cdots & \boldsymbol{\Phi_{1k}} \\ \boldsymbol{\Phi_{21}} & \boldsymbol{\Phi_{22}} & \cdots & \boldsymbol{\Phi_{2k}} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Phi_{t1}} & \boldsymbol{\Phi_{t2}} & \cdots & \boldsymbol{\phi_{15k}} \end{bmatrix}$$

where $\boldsymbol{\Psi}$ is a (*t x t*) diagonal matrix of specific variances. The approach is described as the linear regression of genotype and G×E on latent covariables (the environmental loadings), with a separate slope for each genotype (the genotype scores). All slopes had a common intercept, as the genotype main effects and G×E were not distinguished.

All models were tested with and without the inclusion of ancestry information which was built from historic PFR field books, a publicly available pedigree database (van Berloo et al. 2007), and a catalogue of world potato varieties (Hils and Pieterse 2005). ASReml (Gilmour et al. 2006) and R (R Development Core Team 2012) were used for data analyses, with the mixed models fitted using ASReml-R (Butler et al. 2009). AIC was used as the test criterion for the various forms of G_0. Variance-covariance models were also compared by simulating the response to selection (Piepho and Möhring 2007). With the assumption that breeding values estimated from the data were the 'true' values, residuals at each simulation round (1000 rounds) were resampled with replacement and added to the fitted values. The simulated data were then re-analysed to provide the best linear unbiased predictors (BLUPs) of genotype effects, and at selection fraction *s*, the top (ranked) *s*100% genotypes based on the simulated BLUPs, were identified. The simulated BLUPs were then replaced with the true BLUPs for the selected group of genotypes. The difference between the true breeding value mean of the selected genotypes and the true mean of the breeding population was considered to be the response to selection. Spearman rank correlations, the percentage concurrence of the top-performing genotypes (from truncation selection) and the simulated response to selection were used to compare the *extended* MET evaluation (which included global + extraneous + local within-trial error variation, as identified in single trial analyses) with a *base* MET evaluation (which simply included extraneous (row/column) error variation only).

Breeding value prediction

Empirical breeding values (EBVs) were obtained from the BLUPs of genotype effects, (e.g. Smith et al. 2005, p.458). As variance components were unknown, empirical breeding values resulted from applying variance components in the mixed model equations that were estimated from the data, thus giving empirical BLUEs (best linear unbiased estimators) of fixed effects and empirical BLUPs of random effects. The 95% confidence intervals of EBVs were calculated from the prediction error variances (PEV), so that:

$$95\%CI_{EBV} = EBV \pm 1.96\sqrt{PEV}$$

where PEV is the prediction error variance obtained from the inverse of the coefficient matrix of the mixed model equations for random plant effects. The PEV can be described as the mean phenotype measurement variation of an individual that is not accounted for by the prediction (Cameron 1997).

5.4 Results

Data

Trial yield data is summarised in Figure 5-1 and Table 10-3 (Appendix IV). In these 21 trials, maximum tuber yield was 124 t ha^{-1} for TTY and 108 t ha^{-1} for MTY, both recorded in 1999 (PK-C2-99A). In the PFR potato breeding trials, it is not uncommon to observe yields greater than 100 t ha^{-1}, as estimated from plot trials. The distributions for TTY, MTY and CHI showed a high level of yield variability across trials. Standard deviations (sd), if not the same, were slightly lower for MTY than for TTY. Experimental variability, expressed as a coefficient of variation (CV) was consistently higher for MTY than for TTY indicating that decrease in mean yield from TTY to MTY was relatively greater than the decrease in standard deviation. CVs showed no association with trial dimensions, i.e. number of rows or number of columns, but there were higher CVs for lower mean yields with negative correlations between CV and TTY (-0.47), MTY (-0.69) and CHI (-0.96).

Single trial analysis and spatial effects

Tables 5-1 to 5-3 shows the fitted fixed effects, variance components and spatial parameter estimates for the preferred (*base* or *extended*) model from single trial analyses for TTY, MTY and CHI. Trials were examined for systematic patterns in the random row and column effects, which are considered part of the extraneous error variation as described by Gilmour et al. (1997), but none was identified. Variability of the estimated additive variance (σ_a^2) component was apparent across trials and particularly for C2 trials; for example, the σ_a^2 of MTY ranged from 58 in 2007 to 281 in 2000. The proportion of additive to phenotypic variance (V_P), a measure of the narrow-sense heritability, was still high in these years and was 0.66 in 2007, and 0.77 and 0.85 in 2012. The lowest heritability for MTY was 0.57 in 2001 for trial PK-EM-01.

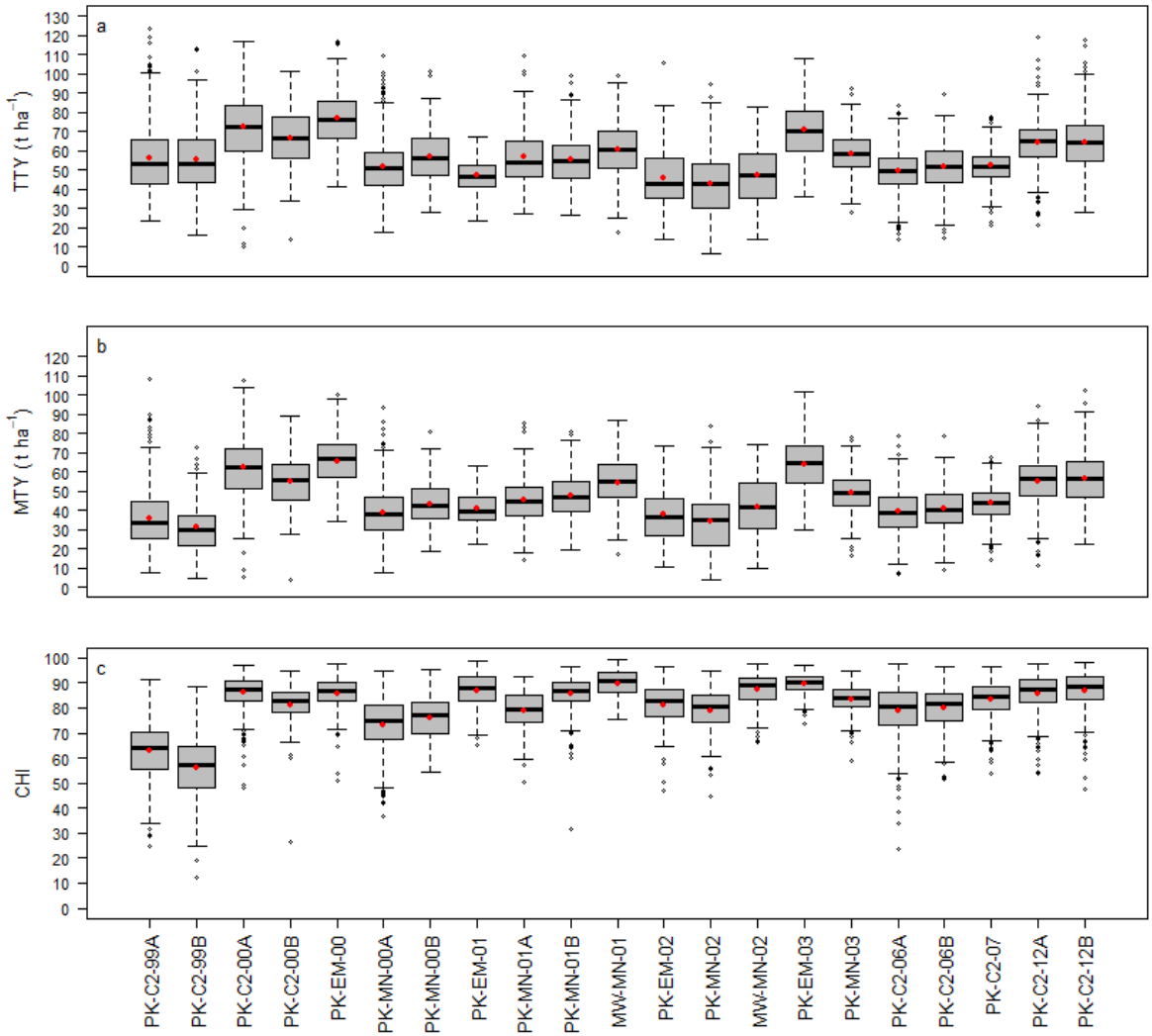

Figure 5–1 Box plots of: (a) total tuber yield (TTY) (tonnes per hectare); (b) marketable tuber yield (MTY) (tonnes per hectare) and; (c) percent marketable yield (untransformed) from 21 early stage potato breeding trials over 8 years. PK and MW trial prefixes refer to Pukekohe and Manawatu locations respectively. Mean yields are indicated by the filled circles

Heritability values, in general, were lower for CHI than for both TTY and MTY (results not shown for TTY and CHI). Fixed linear regressions, in either rows or columns or both, were included for over half of all trials. A second-order polynomial regression (rows) was chosen for trial PK-MN-01B for both TTY and MTY, which appeared to account for curvature present across the trial, as observed from the variogram of residuals. Random spline (row) effects were found to be important in only one trial (MW-MN-02) for TTY and MTY, but there was very little effect on the re-ranking of genotype breeding values in this case. For TTY, the percentage of common genotypes selected from both models was greater than 80 for all trials with the exception of PK-C2-99A (74%). The range of percentage

concurrence of genotypes selected across all 21 trials was greater for both MTY and CHI than for TTY, and included ranges of between 67 and 100 and 66 and 100 for MTY and CHI respectively. AR1 correlation coefficients, when considered to improve model fit, were generally small overall, i.e. mostly between -0.35 and +0.35, but featured more for rows (the shorter plot to plot distance) than for columns, in general. These spatial correlation estimates were all negative for TTY and also for MTY with the exception of two trials (PK-C2-99B and PK-MN-00A). For CHI, there were approximately equal numbers of positive and negative spatial correlations. Because of frequent convergence problems, the 'nugget' effect was not included in analyses. In exploratory analyses of TTY and MTY, examination of a null model (*additive + spatial*) sometimes yielded positive AR1 spatial correlation estimates but these were often effectively reduced to zero or negative when random row and column terms were added. For MTY in trial PK-C2-99A, relatively strong AR1 estimates for rows and columns of 0.46 and 0.42 respectively were reduced to 0.11 and 0 when random row and column terms are added. Similarly for TTY, AR1 parameter estimates for row and columns respectively decreased from 0.37 and 0.41 to -0.07 and -0.21. For this particular trial, the best-fitting model for both MTY and TTY did not include a spatial component. For trial PK-MN-01B, AR1 parameter estimates of 0.28 (row) and 0.19 (column) for MTY were reduced to -0.14 and -0.21 respectively, and for TTY, AR1 estimates of 0.03 (row) and 0.29 (column) in trial PK-C2-07 were reduced to -0.22 and 0.02 respectively, when row and column random terms were added. Spatial effects, when fitted, often had an impact on the concurrence between *base* and *extended* models for the top 10 percent of genotypes selected, but agreement was usually greater than 80%.

Table 5–1 Trial dimensions and REML estimates of variance components from models for total tuber yield (TTY, t ha^{-1}); fixed and random effects and autoregressive (AR1) parameters, Spearman rank correlations (*rho*) and % concurrence of the top 10% of genotypes (ranked on empirical breeding values (EBVs) of TTY) between *base* and *extended* models

Trial code	Dimension	Fixed regressions		Error model				Spatial correlation			*rho*	% concurrence	$\frac{V_A}{V_P}$
			σ_a^2	σ_b^2	σ_r^2	σ_c^2	spl_{row}	ρ_r	ρ_c	σ_e^2			
PK-C2-99A	40×16	l_{row}	l_{col}	326.1		5.4	11.1				38.7	0.94	74	0.86
PK-C2-99B	38×8	l_{row}	l_{col}	248.6		12.3	1.3				56.5	0.98	87	0.78
PK-C2-00A	60×8			256.0		11.6			-0.17	-0.13	76.4	0.99	100	0.74
PK-C2-00B	14×8			358.7		13.1			-0.27		34.7	0.99	83	0.88
PK-EM-00	33×8			150.6	22.3	5.5	2.4		-0.33	-0.14	44.8	0.97	89	0.67
PK-MN-00A	20×24			221.4	23.9	0.9	0.2				61.1	0.94	83	0.72
PK-MN-00B	14×8	l_{row}		192.7		1.3	6.7		-0.27		54.0	0.98	83	0.76
PK-EM-01	24×8	l_{row}		42.3			1.0				20.3	0.99	83	0.67
PK-MN-01A	12×8	l_{row}	l_{col}	256.2					-0.14	-0.23	48.0	0.99	100	0.84
PK-MN-01B	34×8	pl_{row}		123.2		7.0	1.1		-0.22	-0.21	28.0	0.98	92	0.77
MW-MN-01	12×8			198.4		0.17	14.2				57.6	‡	‡	0.73
PK-EM-02	24×12			230.9		3.2	2.2				39.8	‡	‡	0.84
PK-MN-02	14×8	l_{row}		226.9					-0.38		89.2	0.97	83	0.72
MW-MN-02	12×10	l_{row}		166.5		9.7	7.5	8.5			28.4	0.99	100	0.75
PK-EM-03	30×8	l_{row}		177.9		1.9	2.3		-0.28	-0.35	57.5	0.98	88	0.74
PK-MN-03	18×8			108.7			11.7				43.5	‡	‡	0.66
PK-C2-06A	54×10	l_{row}		104.2			11.6		-0.33		56.3	0.92	81	0.61
PK-C2-06B	26×10		l_{col}	140.2	16.4	7.5	5.7		-0.22	-0.21	34.3	0.97	83	0.69
PK-C2-07	34×20	l_{row}		63.4		5.9	5.1		-0.22		16.9	0.99	97	0.69
PK-C2-12A	56×7			154.2		5.5	4.1		-0.32	-0.23	53.4	0.98	89	0.71
PK-C2-12B	54×7			162.3		0.54	3.97		-0.22		93.8	0.99	84	0.62

l_{row} and l_{col} represents a linear regression of the total yield on column or row number respectively; pl_{row} represents a polynomial regression (of order *2*); σ_a^2 is the additive genetic variance; $\sigma_b^2, \sigma_r^2, \sigma_c^2$ and spl_{row} are the subtrial, row and column variances and random row splines respectively and represent the error model; ρ_r and ρ_c are the spatial correlation parameters; σ_e^2 is the residual error variance; V_A/V_P is the proportion of additive genetic variance to the phenotypic variance, a measure of the narrow-sense heritability. ‡ indicates that the *base* model (no spatial error component or trend term) was the best fitting model.

Table 5–2 REML estimates of variance components from models for marketable tuber yield (MTY, t ha^{-1}); fixed and random effects and autoregressive (AR1) parameters, Spearman rank correlations (*rho*) and % concurrence of the top 10% of genotypes (ranked on empirical breeding values (EBVs) of MTY) between *base* and *extended* models

Trial code	Fixed regressions			Error model				Spatial correlation			*rho*	% concurrence	$\frac{V_A}{V_P}$
			σ_a^2	σ_b^2	σ_r^2	σ_c^2	spl_{row}	ρ_r	ρ_c	σ_e^2			
PK-C2-99A	l_{row}		201.0		9.1	12.0				34.7	0.95	81	0.78
PK-C2-99B	l_{row}	l_{col}	121.3			2.9			0.25	43.7	0.93	73	0.72
PK-C2-00A	l_{row}		241.9		4.5					66.1	0.98	92	0.77
PK-C2-00B			280.8		2.2	2.5		-0.41		44.4	0.99	83	0.85
PK-EM-00		l_{col}	108.5	20.9	5.4	1.9		-0.20		37.8	0.97	67	0.62
PK-MN-00A	l_{row}		172.4	25.9	0.9			0.12	0.10	59.2	0.92	79	0.67
PK-MN-00B			143.3		5.2	2.3				44.4	‡	‡	0.73
PK-EM-01	l_{row}		34.2		0.8	1.8			-0.26	23.2	0.97	83	0.57
PK-MN-01A	l_{row}	l_{col}	182.9							46.5	0.98	80	0.80
PK-MN-01B	pl_{row}		94.8		1.3	2.6		-0.14	-0.21	31.5	0.97	77	0.73
MW-MN-01			150.5		0.90	13.4				58.8	‡	‡	0.67
PK-EM-02	l_{row}		181.5		2.3	1.4				37.2	0.99	100	0.82
PK-MN-02	l_{row}		171.7					-0.30		65.5	0.98	67	0.72
MW-MN-02	l_{row}		151.9		11.7	8.3	5.2			26.4	0.99	100	0.75
PK-EM-03	l_{row}	l_{col}	160.6		1.6	1.4		-0.21	-0.31	53.1	0.98	88	0.74
PK-MN-03			94.4			12.1		-0.21		39.4	0.99	80	0.65
PK-C2-06A	l_{row}		109.2		0.5	14.3		-0.30		50.3	0.95	85	0.63
PK-C2-06B		l_{col}	122.2		6.3	3.1		-0.27	-0.29	34.1	0.98	83	0.74
PK-C2-07	l_{row}		58.0		5.2	2.9				21.3	0.99	88	0.66
PK-C2-12A		l_{col}	149.9		4.9	4.4		-0.37	-0.14	48.2	0.98	95	0.72
PK-C2-12B			154.8		0.7	2.2				87.5	‡	‡	0.63

l_{row} and l_{col} represents a linear regression of the marketable yield on column or row number respectively; pl_{row} represents a polynomial regression (of order *2*); σ_a^2 is the additive genetic variance; $\sigma_b^2, \sigma_r^2, \sigma_c^2$ and spl_{row} are the subtrial, row and column variances and random row splines respectively and represent the error model; ρ_r and ρ_c are the spatial correlation parameters; σ_e^2 is the residual error variance; V_A/V_P is the proportion of additive genetic variance to the phenotypic variance, a measure of the narrow-sense heritability. ‡ indicates that the *base* model (no spatial error component or trend term) was the best fitting model.

Table 5–3 REML estimates of variance components from models for % marketable yield (CHI, logit transformed) ; fixed and random effects and autoregressive (AR1) parameters, Spearman rank correlations (*rho*) and % concurrence of the top 10% of genotypes (ranked on empirical breeding values (EBV) of CHI) between the *base* and *extended* models

Trial code	Fixed regressions		Error model				Spatial correlation			*rho*	% concurrence	$\frac{V_A}{V_P}$
			σ_a^2	σ_b^2	σ_r^2	σ_c^2	spl_{row}	ρ_r	ρ_c	σ_e^2			
PK-C2-99A	l_{row}		0.12	0.01	0.02	0.02				0.12	0.97	71	0.41
PK-C2-99B	l_{row}	l_{col}	0.18		0.01	0.02				0.13	0.96	80	0.53
PK-C2-00A	l_{row}		0.16	0.02		0.01				0.16	0.89	75	0.46
PK-C2-00B			0.13		0.01	0.03		-0.34	-0.23	0.15	0.83	92	0.41
PK-EM-00			0.16			0.02			0.20	0.17	0.99	89	0.46
PK-MN-00A			0.11		0.01	0.001		0.21	0.18	0.20	0.97	83	0.34
PK-MN-00B			0.24		0.002	0.005				0.10	0.99	83	0.69
PK-EM-01			0.13		0.01	0.05				0.19	‡	‡	0.34
PK-MN-01A	l_{row}	l_{col}	0.07		0.007			0.25	0.22	0.12	0.96	80	0.36
PK-MN-01B			0.11			0.001		0.30		0.22	0.97	92	0.33
MW-MN-01	l_{row}	l_{col}	0.27			0.01				0.29	0.98	100	0.47
PK-EM-02			0.12		0.007	0.01		-0.11	0.19	0.21	0.99	83	0.35
PK-MN-02			0.09			0.01		-0.13		0.23	0.99	100	0.27
MW-MN-02	l_{row}		0.13			0.05				0.20	0.98	100	0.34
PK-EM-03		l_{col}	0.12		0.002	0.02				0.11	0.99	100	0.48
PK-MN-03	l_{row}		0.08		0.01	0.01		-0.18	-0.39	0.09	0.95	80	0.42
PK-C2-06A			0.27		0.002	0.05				0.15	‡	‡	0.57
PK-C2-06B	l_{row}	l_{col}	0.17		0.01	0.002				0.11	0.96	67	0.58
PK-C2-07	l_{row}	l_{col}	0.11		0.006	0.011				0.15	0.94	66	0.40
PK-C2-12A	l_{col}		0.19		0.009	0.02				0.17	0.97	95	0.49
PK-C2-12B	l_{col}		0.21			0.04		0.17	0.12	0.24	0.98	89	0.43

l_{row} and l_{col} represents a linear regression of the percent marketable yield on column or row number respectively; σ_a^2 is the additive genetic variance; $\sigma_b^2, \sigma_r^2, \sigma_c^2$ and spl_{row} are the subtrial, row and column variances and random row splines respectively and represent the error model; ρ_r and ρ_c are the spatial correlation parameters; σ_e^2 is the residual error variance; V_A/V_P is the proportion of additive genetic variance to the phenotypic variance, a measure of the narrow-sense heritability. ‡ indicates that the *base* model (no spatial error component or trend term) was the best fitting model.

MET analysis and variance models

Table 5-4 shows the highly unbalanced feature of the early-stage trials and the rapid attrition of breeding lines over consecutive years, which is typical of potato breeding programmes. Based on AIC, there was a large improvement of CORH and FA*k* variance structures over DIAG for all three traits. For TTY and CHI, CORH was a small improvement over both FA1 and FA2, with and without the pedigree fitted. The trial-to-trial

Table 5–4 Concurrence of genotypes across 5 years of potato tuber yield trials; diagonal entries are the number of genotypes tested in individual years

	1999	2000	2001	2002	2003
1999	462	114	33	19	8
2000		577	158	64	26
2001			233	101	55
2002				131	68
2003					89

genetic correlation estimate from the CORH model was 0.69 and 0.72 for TTY and CHI respectively. For MTY (with pedigree included) there was no difference in model fit between FA1 and CORH, and for FA2 there was very little improvement over both FA1 and CORH. For MTY, the trial-to-trial genetic correlation estimate from the CORH model was 0.69. A heatmap plot of REML estimates of the genetic correlations from the FA model for MTY is shown in Figure 5-2. There was a pattern of decreasing genetic correlations over time, which may have partly reflected the limited concurrence of genotypes in trials two or more years apart but there were no negative genetic correlations estimated between any trials. For MTY, the lowest genetic correlations were found between PK-MN-03 and the 1999 to 2000 trials. These ranged from 0.06 to 0.39. For MTY, fitting a FA1 model accounted for 71% of the variance, which increased to 77% with a FA2 fit. With regard to individual trials, the FA2 model provided a reasonable fit with the exception of trials (in ascending order) PK-MN-02, MW-MN-01, MW-MN-02 and PK-EM-02 with 36, 42, 52 and 53% of variance accounted for respectively, by the two latent variables. This indicated that these trials were, in general, less well correlated with the other trials. It is also interesting to note that the second latent variable (data not shown) was a temporal contrast between 1999 and 2000 trials (that included the C2 stage trials), and 2001, 2002 and 2003

trials (with the exception of PK-MN-01, which was in the 1999 to 2000 grouping). There were problems with convergence for the FA3 models, as there were with attempts to fit an unstructured (US) model for all three traits. Including a relationship matrix (pedigree) in the analyses improved model fit for TTY and MTY but not for CHI (Table 5-5).

Table 5–5 Summary of genetic variance models (G_0), number of genetic and total variance parameters (N_p) and Akaike information criterion (AIC) goodness-of-fit for total tuber yield (TTY), marketable tuber yield (MTY) and the fraction of marketable yield (CHI)

			TTY		**MTY**		**CHI**	
	N_p		AIC		AIC		AIC	
G_0 structure	G_0	total	$\boldsymbol{I}\sigma_g^2$	$\boldsymbol{A}\sigma_a^2$	$\boldsymbol{I}\sigma_g^2$	$\boldsymbol{A}\sigma_a^2$	$\boldsymbol{I}\sigma_g^2$	$\boldsymbol{A}\sigma_a^2$
SIMPLE	1	54	405	311	345	285	98	121
DIAG	15	68	362	216	298	241	113	135
CORH	16	69	115	0	49	6	0	6
FA1	30	83	126	8	81	6	7	16
FA2	44	97	121	2	82	0	22	31

AIC expressed as the difference from the best fitting model. $\boldsymbol{I}\sigma_g^2$ represents the independent genotypic variance (no pedigree fitted) and $\boldsymbol{A}\sigma_a^2$ represents the pedigree-based genotypic variance (pedigree included)

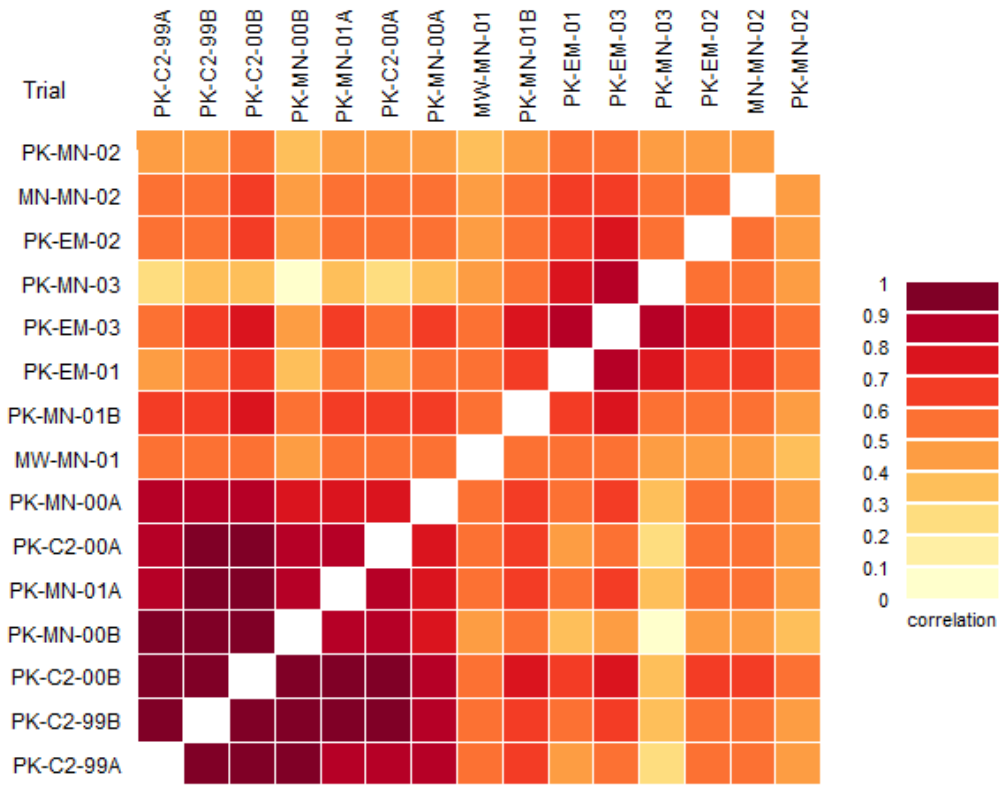

Figure 5–2 Genetic (additive) correlation estimates of MTY (marketable yield) between early-stage potato selection trials from a factor analytic structure of order 2 (FA2)

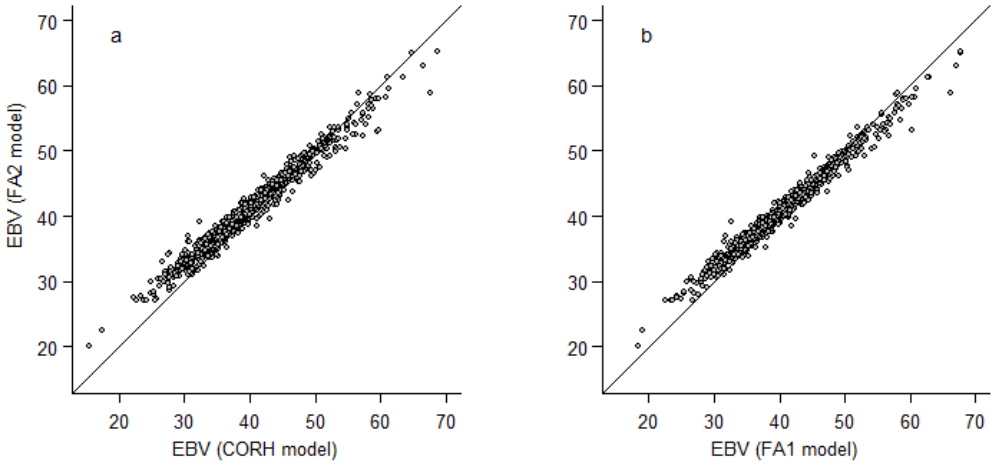

Figure 5–3 Scatterplots of empirical breeding values (EBV) for marketable tuber yield (MTY) for: (a) EBVs predicted from heterogeneous variance-homogeneous correlation model (CORH) and factor analytic structure FA2 models and; (b) EBVs predicted from FA1 and FA2 models

Empirical breeding values for tested genotypes predicted from CORH, FA1 and FA2 were all highly correlated with each other, with product-moment correlation coefficients between 0.98 and 0.99. There was a 93% concurrence of the top ranked 10% of selected genotypes between FA2 and FA1 and 91% between FA2 and CORH. Figure 5-3 illustrates a shrinkage effect of MTY when data were fitted to a FA2 model, with plots of EBVs slightly departing from the line of unity between: (a) CORH and FA2, and (b) FA1 and FA2. Simulations of the response to selection over all levels of *p* (the proportion of the top ranked genotypes selected) were similar for CORH and FA1 but were reduced for FA2, which also reflected the shrinkage of empirical breeding values (Fig. 5-4).

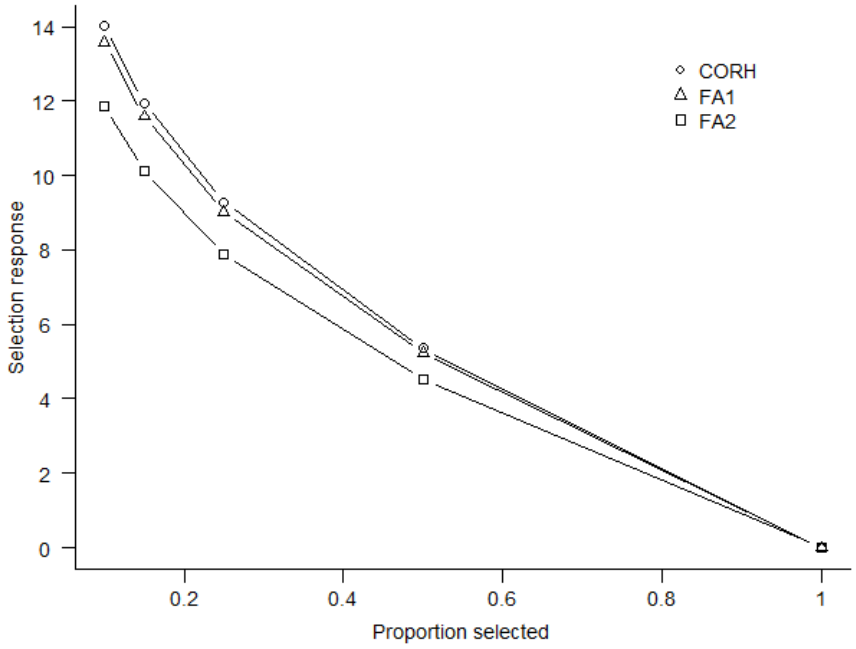

Figure 5–4 Simulated response to selection for marketable tuber yield (MTY, t ha^{-1}) for variance models heterogeneous variance-homogeneous correlation model (CORH), factor analytic structure FA1 and FA2

In simulations to compare *base* and *extended* MET evaluation models, Spearman rank correlations were high (0.98 to 0.99) and genotype concurrence for the top ranked 10% of selections was 0.93, 0.92 and 0.91 for CORH, FA1 and FA2 respectively. The simulated responses to selection for the *base* MET models (for variance models CORH, FA1 and FA2) were very similar to those simulated for the *extended* MET models, and each was within the 95% (parametric) bootstrap confidence interval obtained for the response to selection from their respective *extended* models.

Empirical breeding values for MTY and their 95% confidence intervals are shown in Table 5-6 for a number of genotypes that were tested between 1999 and 2003, including some internationally recognised cultivars that were used as parents in the PFR breeding programme.

Table 5–6 New Zealand Empirical Breeding Values (EBV) and their 95% confidence intervals (95% CI_{EBV}) for marketable tuber yield (MTY); examples of cultivars and advanced clones used as parents of 1999–2003 tested genotypes

Variety	†EBV	95% CI_{EBV}	Female parent	Male parent	‡Year
'Summer Delight'	20.4	8.0	1858.21	V394	1990
'Moonlight'	17.0	5.9	1463.1	V394	1987
'Pacific'	10.9	8.2	'Tekau'	V394	1981
'Spunta'	9.6	15.6	'Bea'	USDA96-56	1968
'Kaimai'	8.5	5.8	'Rua'	V394	1992
'Karaka'	7.0	4.2	002/9	V394	1992
'Kennebec'	6.4	9.8	USDAB127	USDA96-56	1948
'Driver'	6.3	6.9	993.6	V394	1983
'Horizon'	6.0	9.4	1053.57	'Baillie'	1991
'Van Gogh'	5.5	11.2	Zpc69-C-239	'Gloria'	1989
'Katahdin'	4.5	11.5	USDA40568	USDA24642	1932
'Fraser'	4.4	9.4	676.34	'Whitu'	1990
'Asterix'	4.2	14.3	'Cardinal'	VE70-9	1991
'Glenna'	4.1	16.2	10223-7	10300-13	1987
L115-1	4.0	5.7	H612-3	D-4	1970
'Agria'	3.8	5.9	'Quarta'	'Semlo'	1985
'White Delight'	2.6	10.0	002/9	'Maris Piper'	1982
'Gladiator'	0.6	5.8	B5281-1	Vtn62-33-3	1995
'Brodick'	-0.8	6.7	7683a-12	8898AC-14	1990
'Fianna'	-1.0	5.3	Konst62-660	AM64-2	1987
V394	-1.2	3.5	D47/11	D42/8	1970
'Ranger Russet'	-1.2	3.8	'Butte'	A6595-3	1991
'Atlantic'	-2.2	7.9	'Wauseon'	'Lenape'	1976
'Coliban'	-3.6	10.4	'Kennebec'	V28-12	1974
'Red Rascal'	-4.0	10.0	'Tekau'	'Desiree'	1981
L118-2	-4.3	7.1	H614-1	D-4	1970
'Nadine'	-6.1	7.8	('Desiree' × 'M. Piper') × vrn-seedling	'Pentland Dell' ×vrn-seedling	1987
'Desiree'	-6.4	10.6	'Urgenta'	'Depesche'	1962

†Expressed as a deviation from the mean EBV of C2 tested genotypes; ‡year of commercial release or first use as parent in New Zealand (approximate)

5.5 Discussion

Spatial parameters and interplot competition

Local spatial trends were not a consistent feature of the potato trials tested; extending models to include spatial effects did not always improve model fit. Relative efficiencies also indicated that there was only a small benefit for including local and/or (global) field trends in the potato trials analysed and occasionally there no benefit at all. Maximum RE (%) was 111, 109 and 109 with means of 102, 101 and 102 for TTY, MTY and CHI respectively. In contrast, a number of studies in other field crops, particularly cereals, have demonstrated consistent advantages when including spatial terms, and relative efficiencies have often shown large improvements. In wheat studies for example, Qiao (2000) found a greater

improvement for long-faced (single column by many rows array) trials than row-column trial arrangements, but obtained an overall mean RE of 138, with 11 out of 33 trials having an RE greater than 140%. Dutkowski et al. (2006) found that the evaluation of forest genetic trials were often enhanced when augmenting models with spatial components but the extent of any improvement was trait-dependent. Spatial analysis has been considered as an alternative model to the traditional analysis of complete or incomplete block designs, but based on the comprehensive re-analysis of 53 lentil variety trials, Sarker et al. (2001) recommended that block design methods could often be enhanced but not replaced with spatial methods. Müller et al. (2010) also emphasised that prudence was probably the best approach and advised that over-complication should be avoided when extending block models with spatial effects. They found that a standard block model outperformed a spatial model in most cases when analyzing 293 sugar beet and 64 barley trials. From the results presented in the current study, the blocking designed for trials often appeared to deal adequately with localised heterogeneity. There was also evidence of non-stationarity, so that fitting fixed linear regression often reduced global trends of residuals, but higher order polynomials or cubic splines were rarely helpful. When global and local spatial trends were neglected in the subsequent MET analysis, there was very little re-ranking of genotypes compared with an extended MET evaluation, suggesting that extending models to take account of these effects may make very little difference to the realised genetic gain of potato yield. That said, the effort expended in checking for spatial effects is small, compared with the effort and costs involved in setting up and managing field trials. Spatial modelling should therefore be a consideration in potato evaluation to account for possible systematic field heterogeneity that may be caused by localised factors within a trial site, such as soil chemical and physical properties (e.g. Redulla et al. 2002; Po et al. 2010).

The absence of strong spatial correlations contrasts with previous results in annual wheat and canola crops (e.g. Gilmour et al. 1997; Oakey et al. 2006; Beeck et al. 2010) and perennials such as grapes (Gonçalves et al. 2007), which have reported high autoregressive parameters of greater than 0.7. Even when spatial effects did appear to be important, spatial correlation estimates were often small and mostly negative (for TTY and MTY) and their magnitude was commonly reduced by subsequently fitting fixed regressions to account for global field trends. Similarly, small negative spatial correlations were also found by Stringer et al. (2002) in sugarcane trials and were attributed to interplot (inter-genotype) competition. The use of larger interplot distances has been suggested as a means to eliminate plot competition in breeders' trials, but this implies a reduction in selection

intensity when the total trial area is fixed. Furthermore, bias due to competitive ability may be replaced by bias due to the occurrence of a genotype × plant density interaction, whereby the performance of genotypes is density-dependent (Bos 1983b; Bos and Caligari 2008). Modelling approaches to account for fertility trends and interplot competition simultaneously have been developed in sugar beet (Durban et al. 2001) and sugarcane (Stringer et al. 2011). Work by Connolly et al. (1993) identified competitive effects of yield in single-row plots of potatoes, although these effects were not ubiquitous over all trials tested. They found little re-ranking of genotypes but there was shrinkage in the range of yield estimates from high and low yielding plots and closer agreement with pure-stand yields after accounting for competitive effects.

Genetic variance models

For the quantitative genetic evaluation of MET yield data in early-stage potato breeding trials, a number of different genetic variance-covariance models were compared. A heterogeneous variance and homogeneous correlation structure (CORH) was found to be adequate for modelling the G×E effects for TTY and CHI in the early-stage trials tested. For MTY, there was only a small difference between the AIC of the factor analytic model of order 2 (FA2) AIC and both the FA1 and CORH models. This suggested that these three models performed equally well for the analysis of MTY. EBV plots between CORH and FA2, and FA1 and FA2 (Fig. 5-3) indicated that the FA2 model was possibly over-fitting the data and some degree of bias was introduced into the process when moving from a FA1 to a FA2 model. This was also reflected in the simulations of the responses to selection in which both CORH and FA1 were similar over all levels of *p*. In comparison, responses were reduced when selections on EBVs were based on simulations from FA2 (Fig. 5-4).These results also suggested that environments, which were mainly temporal in the current study, i.e. different years or growing seasons in the same location, were relatively homogeneous for these data. The results are, of course, presented in the context of early-stage trials under New Zealand conditions when the extent of MET testing is limited. From model comparisons, other studies have found that FA variance structures were suitable for both early- and late-stage evaluation trials (Smith et al. 2001; Crossa et al. 2006; Kelly et al. 2007; Burgueño et al. 2011). These studies were usually based on MET data from more extensive trial data and, most probably, more diverse environments, such as the extensive international wheat trials from the study of Crossa et al. (2006). Where access to suitable software and knowledge is limited, evaluation by fitting a homogeneous correlation

structure may provide a more approachable and parsimonious method to fit a genetic variance structure to MET data. From cross-validation studies, Burgueño et al. (2011) found that when G×E was not complex, i.e. a re-scaling of performance rather than a re-ranking of genotypes, both FA and simple non-FA (ignoring G×E) models gave good predictive ability. So and Edwards (2011) found that because of poor genetic links of maize hybrids across environments, modelling heterogeneous genotype covariances did not improve predictions. For potato, there are situations where a simple homogeneous correlation model may be expected to be less suitable, such as the analysis of yield data from multiple locations that represent a greater diversity of environments. As selection stages progress in the PFR programme, METs are expanded to more locations throughout New Zealand, but these may still be relatively homogeneous compared with results from other programmes that test genotypes across a greater range of latitudes.

Use of ancestry information

For pedigree-based BLUP of breeding values in plants and animals, it is recommended that all data that have been used in selection decisions should be included in the evaluation for the estimate of breeding values (Piepho and Möhring 2006). In the present study, different variance models were tested and breeding values estimated from C2, C3 and C4 data only. Therefore, an assumption of this analysis was that selection for yield was absent in the initial generations (seedling and C1). This seemed reasonable for TTY and MTY, as there had been no explicit selection of yield in these generations. There is no explicit selection for yield in the seedling or C1 generations, as previous work has considered the low efficiency of potato selection in the early generations, showing a poor association between selection for performance as seedlings and performance in the C1 generation (Brown et al. 1984; Brown and Caligari 1986). Further work also demonstrated that selection for yield at the first clonal stage was "*only marginally more effective than a random reduction in number of genotypes*" (Caligari et al. 1986). CHI, however, was highly correlated with the general impression score, which is a categorical preference score of tubers given on a 0 to 9 scale by breeders, in the present study (results not shown). General impression is a trait for which there has been selection in the initial generations but the basis on which these decisions were made was not recorded. Using a pedigree-based genotypic variance, V_G (i.e., $V_G=\mathbf{A}\sigma_A^2$), as in the present study, Piepho and Möhring (2007) reported a better model fit in some analyses when assuming independence between genotypes (i.e., $V_G=\mathbf{I}\sigma_G^2$). It was suggested that selection has possibly taken place and the information on which selection

had been based was not included in the analysis. A cautious approach may therefore be prudent and involve the testing of both pedigree and non-pedigree based models. The additive relationship matrix **A** used in the current study was based on disomic inheritance. Under the assumptions of no past selection, double reduction or inbreeding, the expected additive genetic covariances both of diploid and tetraploid relatives is equivalent (Lynch and Walsh 1998), which may or may not be appropriate when dealing with autotetraploid potato. From our, as yet unpublished, analysis of other potato data, there was very little difference in the BLUPs when a diploid relationship matrix was replaced with a tetrasomic-based relationship matrix that was derived using the method of Kerr et al. (2012). Oakey et al. (2007) partitioned the non-additive genetic effects into dominance and residual genetic components. Fitting a non-additive component may reduce bias in breeding value estimation and variance estimates could be exploited by selecting favourable parental combinations (Mrode 2005; Oakey et al. 2007). It may also be more appropriate for clonal selection in cultivar development i.e. selection of individuals with a high total genetic value – a high 'potential production ability' or 'production worth'. This demands further investigation in potato, but satisfactory estimates of non-additive genetic effects require large breeding populations (e.g. Pante et al. 2002).

Empirical breeding values for potato tuber yield

Marketable yield, as defined in this particular study, is a rather generic definition and may not, in itself, fully describe a genotype's economic worth in terms of tuber yield. Payment schedules written into contracts for processing potatoes, for example, are sometimes based on tuber size or weight, so that there are penalties or premiums for tubers that fall within particular thresholds. North American programmes often specify a 'US No.1 Yield' which is the tuber yield criterion that maximises the economic yield objective for French fry processing. Tuber size (or weight) distribution is therefore a factor in determining farm revenue and suggests that such a measure will go some way to better describe the economic worth of genotypes and this may vary depending on the target end-use. The cultivar 'Summer Delight' for example, was shown to have a high New Zealand EBV for MTY in this study (Table 5-6) but is known to produce a high proportion of large tubers; tubers that are considered too large for the New Zealand table (fresh) market sector but are a suitable size for factory (French fry) processing. Tuber size can, to some degree, be manipulated by closer plant spacing but this will result in increased seed tuber costs. The definition of MTY in its present form is, therefore, fairly crude and is over-simplistic for a programme that

develops cultivars for several industry sectors, suggesting the need for customized yield indices depending on the specific end-use.

Approaches to increase rates of genetic gain are ongoing areas of research in plant breeding, with selection methods now being augmented with molecular information in a number of crops. These methods target specific areas of the genome that explain a large proportion of trait variation (e.g. Schultz et al. 2012), or whole-genome evaluation that provides predicted genomic breeding values for selection (e.g. Heffner et al. 2009; Lorenz et al. 2011), which is promising for providing the best prospects for increasing the rate of genetic gain in potato tuber yield. Under the resource constraints of modern breeding programmes there is growing interest in estimating breeding values with dense, genome-wide markers (Lorenz et al. 2011), as such techniques may increase selection intensities and reduce both generation intervals and testing requirements. Using this approach, more detailed measurements, such as mean tuber weight and tuber size distribution, would be beneficial but only practically feasible using automatic and rapid phenotyping and data collation techniques. The accumulation of such data would assist the move towards molecular breeding technologies and help to gain more insight into the genetic control of yield and its determinants which, in turn, would help molecular geneticists, agronomists and crop modellers as well as plant breeders. Accurate interpretation of molecular data for the prediction of genomic breeding values will, however, rely upon precise estimation of genotype effects from the phenotypic observations of field 'training' populations. These will, in turn, rely upon the removal of environmental effects and causes of biased estimates, such as the non-independence of plot errors and interference, and a greater understanding and appreciation of the G×E terms in potato breeding trials.

5.6 Conclusions

The use of historic field data provides an opportunity to explore statistical models that improve the methods and precision of identifying new high-yielding genotypes for use as parents, as well as potential and worthy cultivars. In the analyses of potato field trials, spatial effects were not important in all years but there was evidence of interplot competition in some years. The fitting of local and global trends often resulted in some changes in the prediction of top-ranked genotypes compared with a baseline row × column error model, but were unlikely to increase the realised rate of genetic gain. For the genetic evaluation of potato yield, a homogeneous correlation structure to model G×E effects

(allowing for heterogeneity of trial variance) was preferred for TTY and CHI for the series of early-stage MET trials tested. There was little difference between the use of a factor analytic model and homogeneous correlation model for MTY. A better understanding of trial heterogeneity in early-stage potato breeding trials should allow breeders to re-evaluate conventional selection strategies and help to improve the molecular-based analyses of traits, using field data routinely collected in breeding programmes.

6 Selection for tuber yield in a potato improvement programme: variety performance and stability evaluation from multi-environment trials

6.1 Summary

Differences in the yield responses of genotypes across environments, or genotype-by-environment (G×E) interaction, hinder the progress of genetic improvement. Characterisation of these effects helps to determine breeding strategies and improve resource allocation in a cultivar development programme. This study used historical multi-environment trial (MET) data for the analysis of 34 trials in five locations for potato marketable yield in a New Zealand potato improvement programme. Using a factor analytic mixed model, contrasts based on the environmental loadings were observed between the programme's main trial locations in the North Island (Pukekohe) and the South Island (Lincoln), indicating that these locations optimised differentiation between genotypes in terms of G×E effects. Genetic correlation estimates between trial environments were mostly moderately high (>0.5) to high (>0.8) and ranged from zero to positive, with a maximum coefficient of 0.97, suggesting that quantitative (re-scaling) rather than qualitative (crossover) G×E interaction effects were of greater importance. A number of newly developed varieties were shown to have higher genetic yield potential than older and established commercial cultivars, but did not necessarily show better yield stability over the locations tested.

6.2 Introduction

In plant breeding, the accurate discrimination between genotypes is confounded by the main genotype effect and the effect of a genotype-by-environment interaction when testing selection candidates over multiple environments (Bos and Caligari 2008). The larger the genotype-by-environment (G×E) interaction component, the greater the reduction in the correlation between phenotypic and genotypic values, therefore increasing the difficulty of identifying superior genotypes and compromising genetic progress from selection (Cooper and DeLacy 1994). A better understanding of G×E effects within a MET (multi-environment trial) testing regime allows a re-evaluation of resource allocation and selection strategy in a breeding programme. The type and extent of G×E is of particular interest to plant breeders as the characterisation of environments will help, in part, to define selection

strategies. For example, measures of quantitative G×E (heterogeneity of variance, or the scale-change of genotypes) between test locations may help to determine that some locations offer little extra information in terms of differentiating genotypes, i.e. a certain degree of 'environment duplication' is present that incurs an opportunity cost. Consideration should therefore be given for such locations to be dropped from the testing schedule. Alternatively, the presence of qualitative G×E (crossover interaction, or the rank-change of genotypes) may determine that separate breeding programmes for subsets of locations are necessary to select for specific adaptation.

There are numerous statistical approaches to model G×E effects in plant breeding, which have been extensively reviewed by various authors (e.g. Fox et al. 1997; van Eeuwijk et al. 2005; Crossa et al. 2010). In general, these methods are based on univariate or multivariate methods that vary in their degree of complexity and the information that they provide. Over recent years, flexible multivariate multiplicative methods have found favour, including the 'additive main effects and multiplicative interaction' (AMMI) model (Gauch and Zobel 1988; Crossa et al. 1991). This approach carries out singular value decomposition on the matrix of the two-way table of G×E effects, whereby each is modelled as the product of a genotypic score and an environmental score (or loading). Additional multiplicative (bilinear) terms are considered if they improve model fit. AMMI is classified as one of several types of general linear-bilinear model. More recently, a multiplicative mixed modelling approach using factor analysis, which is considered as another class of linear-bilinear model and a mixed model analogy of the AMMI fixed-effect model (Piepho 1997, 1998; Smith et al. 2001; Smith et al. 2005), has been used to evaluate MET data. Heavy attrition of breeding lines at each stage of a MET series of breeding trials is typical of plant breeding programmes and the incomplete nature of such data is better dealt with by residual maximum likelihood (REML)-based procedures. Further, there has been growing trend amongst crop breeders, following their animal and tree breeding counterparts, to treat genotypes as random effects, at least in the early stages of trials. Genotype value predictions are 'shrunk' towards the mean to allow for the uncertainty surrounding the distribution of random effects and there is greater flexibility in analyses with, for example, inclusion of a coefficient of coancestry matrix to take account of relationships among genotypes (Crossa et al. 2006; Oakey et al. 2006; Piepho et al. 2008a).

Potato crops are known to show variability in seasonal yields over both regional and field scales (Redulla et al. 2002; Po et al. 2010), suggesting the need for extensive MET evaluation. Studies into G×E interactions in potato breeding studies have generally been

restricted to a limited number of advanced clones and cultivars, (e.g. Tai and Coleman 1999; Cotes et al. 2002; Affleck et al. 2008). Typically, breeders require information on a larger number of genotypes for inference of performance and stability to help to make more informed selections earlier in a breeding programme, and there is a desire to distribute clones across multiple location as early as possible (Haynes et al. 2012a). The potato breeding programme of The New Zealand Institute for Plant & Food Research Limited (PFR) targets the selection of genotypes that perform well across all major potato-production regions i.e. those that are broadly adapted within a New Zealand context. The multivariate analysis of MET data provides an opportunity to assess the extent and type of G×E present in historic potato yield trials, which may go some way to guide resource allocation for METs and the testing strategy for genotype selection in future, by evaluating environments as well as genotypes.

The study takes a mixed linear-bilinear modelling approach to measure G×E effects and stability of genotypic responses across the major potato production regions of New Zealand for potato yield. It uses data collected from a series of historic yield trials over a number of year-location combinations (environments). A factor analytic (FA) model is used to measure the relationships between genotype performance and environments, and to characterise environments. The yield performance and the stability of recent advanced selections from the PFR breeding programme are also compared with those of older established cultivars that are currently and in some cases widely grown in New Zealand. The study aims to evaluate test locations that are used for the selection of broadly adapted cultivars to improve selection efficiency. Potato varieties are also evaluated to assess the genetic progress of tuber yield improvement in the New Zealand potato breeding programme.

6.3 Materials and Methods

Data

The data for study were collected from breeding trials of the PFR potato genetic improvement programme. Trials were performed between the years 1999 and 2005 at five sites that represent the major potato-producing regions in New Zealand (Fig. 6-1): Pukekohe, South Auckland (37° 12'S 174° 57'E, 141 m above sea level (asl)); Matamata, Waikato (37° 48'S 175° 45'E, 53 m asl); Ohakune , Central (39° 24' 175° 24'E, 741 m asl), Palmerston North, Manawatu (40° 21'S 175° 37'E, 30 m asl) and Lincoln, Canterbury (43°

39'S 172° 28'E, 14 m asl). Pukekohe and Lincoln are the main potato research sites and along with Manawatu, can be categorized as PFR 'on-station' trials, as trials are all managed on PFR research farms. Waikato and Ohakune are 'off-station' trials as these are managed within a commercial potato crop. The 34 trials, synonymous with 'environments' (year-location combinations), were clonal stages four and five (C4 and C5 respectively) of 'main' crop tuber yield trials. Target planting dates, harvest dates and the number of genotypes entered into each trial are shown in Table 6-1. C4 trials were only carried out in Pukekohe (PK), Palmerston North (MW) and Lincoln (LIN). Selected genotypes from the C4 stage were entered into further main crop trials at stage C5 that also included locations Waikato (WAI) and Ohakune (OHA) as well as PK, MW and LIN. Each season in the Waikato region, there were two trials: an early (E) trial and a late (L) trial. The late trial represented the regional practice of the winter-harvesting of potatoes so that the crop is maintained in the ground for approximately 200 days.

Table 6–1 Summary of target potato planting and harvest dates (1999 – 2005), canopy days, days from planting to harvest and number of lines tested per trial

Trial location	Planting date	†Canopy days	Harvest date	Days-to-harvest	Clones/ trial
Pukekohe (PK)	1 November	140	20 May	200	40 – 60
Manawatu (MW)	25 October	140	10 April	170	35 – 55
Ohakune (OHA)	10 November	140	1 June	200	20
Waikato E/L (WAI)	1 Oct./10 Nov.	120/140	1 March/1 June	150/200	24/24
Lincoln (LIN)	10 October	130	10 April	180	60

All dates and days are approximate. Waikato E/L refers to Waikato early and Waikato late trials, respectively. †Canopy days are the number of days from planting to canopy loss (by natural senescence, desiccation (by chemical means) or mechanical destruction).

Trials at all North Island locations (PK, MW, WAI, OHA) were based on Latinized row-column designs typically of 20 to 60 genotypes replicated twice (C4 trials) or three times (C5 trials). Each genotype occurred once, at most, in both rows and columns across a trial of rectangular plots. A typical plot was made up of 12 tubers in total, planted in a six by two arrangement. The South Island trials (LIN) were randomised completed block (RCB) designs, replicated three times. Plots were made up of 12 tubers in total, planted in a 12 by one arrangement.

Each plot was harvested and yield was recorded as marketable tuber yield after undersized (less than 80 g) and defective tubers had been removed. Defective tubers, for example, may have secondary or abnormal growths, rot or excessive greening. Plot yield

was converted to t ha^{-1} (metric tonnes per hectare) for analyses of marketable tuber yield. Marketable yield, as described and hereon in referred to simply as 'tuber yield', is usually considered to be the total economic yield as there is often no economic value attributed to undersized, but otherwise sound, tubers. Although there were a total of 1619 genotypes represented in the data, many were lost after only two years of testing through the discarding of unsuitable candidates. Genotypes of particular interest were tested in at least four locations over three years, and were made up of both New Zealand and international cultivars as well as advanced clonal selections. Many of the international cultivars are popular commercial cultivars widely grown for fresh and processing production in New Zealand.

For clarification, the term 'variety' is used generically and can refer to both clonal selections and cultivars. The term 'cultivar' is used to describe a variety that has been named and commercially released and is protected under Plant Variety Rights (PVR).

Figure 6–1 Location of PFR potato breeding MET trials representing the major potato producing regions in New Zealand. Main research sites (▲) are at Pukekohe and Lincoln

Statistical model

For illustration, the general form of the linear mixed model for the j^{th} trial (environment) was: $\mathbf{y_j} = \mathbf{X_j m_j} + \mathbf{Z_{1_j} b_j} + \mathbf{Z_{2_j} g_j} + \mathbf{e_j}$ where $\mathbf{y_j}$ is the vector of yield observations, m_i denotes the fixed effects of trial means, $\mathbf{b_j} \sim N(0, \mathbf{I}\sigma^2_{b_j})$ and $\mathbf{g_j} \sim N(0, \mathbf{I}\sigma^2_{g_j})$ are vectors of random (non-genetic) design factors, and genetic effects respectively, and $\mathbf{e_j} \sim N(0, \mathbf{I}\sigma^2_{e_j})$ is the vector of random error terms. $\mathbf{Z_{1_j}}$ and $\mathbf{Z_{2_j}}$ are known incidence matrices of 0s and 1s that relate the phenotypic observations to their corresponding vectors and **I** are identity matrices. The non-genetic factors were trial blocking factors and included the rows and columns of the incomplete block designs (PK, MW, WAI, OHA) and the complete blocks of the RCB designs (LIN). For MET analysis, the mixed model equations (MME) were constructed to analyse the vector of observations for the 34 trials $\mathbf{(y_1, y_2, \ldots, y_{34})}$ tested from 1999 to 2005. The joint distribution of the random terms was assumed to follow a multivariate normal distribution with means and (co)variances defined by:

$$\begin{pmatrix} \mathbf{b} \\ \mathbf{a} \\ \mathbf{e} \end{pmatrix} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{bmatrix} \mathbf{B_0} \otimes \mathbf{I_b} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G_0} \otimes \mathbf{I_g} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \otimes \mathbf{I_e} \end{bmatrix} \right]$$

where **0** are null matrices. $\mathbf{B_0}$, $\mathbf{G_0}$ and **R** are covariance matrices for design factors (row and column, or block), genetic and residual effects, respectively, and $\otimes$ is the direct (Kronecker) product. The matrix $\mathbf{B_0}$ is a diagonal matrix of (non-genetic) scaled identity matrices and plot error effects **R** are assumed to be block diagonal and **I** are identity matrices. The assumption was that the variance matrix for the genotype effects has the separable form $\mathbf{G_g} = \mathbf{G_0} \otimes \mathbf{I}$, where $\mathbf{G_0}$ is the matrix of genetic variances and covariances between environments and **I** is an identity matrix. **I** can be replaced by **A**, with the coefficient of coancestry between genotypes as elements – the numerator relationship matrix (e.g. Crossa et al. 2006; Kelly et al. 2009).

Three forms of the genetic variance matrix, G_0, were first tested against each other to complement work on fitting genetic variance structures to potato MET data in Chapter 4. With diagonal elements representing genetic variances for each trial (environment) and the off-diagonal elements representing genetic covariances between pairs of trials, definitions of the tested forms of G_0 were as follows:

DIAG: within-year variances are assumed to be different and all pairwise covariances between (*j*) trials are assumed to be zero

CORH: variances within trials are assumed to be different and a single correlation is assumed between all pairwise combinations of *j* trials.

FA*k*: factor analytic (Piepho 1998; Smith et al. 2001), which is applied to the G×E table of phenotype means and described as the linear regression of genotype and G×E on latent covariables (the environmental loadings) with a separate slope for each genotype (the genotype scores) and separate or common intercept(s), depending on whether genotype main effects and G×E are combined or fitted separately. There was no distinction between genotype main effects and G×E effects in the present analysis. The loadings (as the leading principal components) and residuals, or specific variances, are given by: $\mathbf{G_0} = \mathbf{\Lambda\Lambda'} + \boldsymbol{\varphi}$; where $\mathbf{\Lambda}$ is a (*j x k*) matrix of environmental loadings (ϕ_{jk}) and $\boldsymbol{\varphi}$ is a (*j x j*) diagonal matrix of specific variances:

$$\mathbf{\Lambda} = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1k} \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{j1} & \Phi_{j2} & \cdots & \Phi_{jk} \end{bmatrix}$$

Therefore, the random effect of genotype *i* in environment *j*, (g_{ij}), as presented by Yang et al. (2009), is a linear function of latent variables x_{ik} with coefficients $\Phi_{jk}$for latent factor *k* = 1,2…*p* (where *p* is usually ≤ 3 for crop evaluation trials), plus the lack of fit, φ_{ij}, so that:

$g_{ij} = \mu_j + \sum_{k=1}^{p} x_{ik}\,\Phi_{jk} + \varphi_{ij}$

with g_{ij} approximated because of the usual retention of the first 3 (or less) important latent factors and the ij^{th} cell mean of the genotype × environment table is $y_{ij} = g_{ij} + e_{ij}$, where y_{ij} is the phenotypic value and e_{ij} is the residual error term of genotype *i* in environment *j*. A rotation is applied to the matrices of genotypic scores and environmental loadings to obtain a principal component solution as a more useful interpretation. For genotypes, when there is no distinction between genotype main effects and G×E effects, the first score factor mainly represents genotype performance and the second score factor illustrates genotype stability (Stefanova and Buirchell 2010) and is defined as static stability, where yield performance is consistent across environments (Annicchiarico 2002).

AIC was used as the test criterion for the various forms of G_0 with the Akaike information criterion (AIC) goodness-of-fit test: AIC = -2(*logl* - V_p) where *logl* is the

REML estimate of the log-likelihood and V_p is a penalty term representing the number of variance parameters fitted. Lower values for AIC represented a better fitting model. The analyses of the data were undertaken using ASReml (Gilmour et al. 2006) and R (R Development Core Team 2012), with the mixed models fitted using ASReml-R (Butler et al. 2009).

The best linear unbiased estimators (BLUEs) of the fixed effects, $\hat{m}$, and best linear unbiased predictors (BLUPs) of the random effects ($\hat{b}$ and $\hat{g}$) were obtained from the solutions to the mixed model equations (MME) (Lynch and Walsh 1998). Variance components are unknown and were estimated from the data. Empirical genetic values are therefore a result of applying variance components in the MME that are estimated from the data, so providing empirical BLUEs of fixed effects and empirical BLUPs of random effects. The 95% confidence intervals of BLUPs of genotype values were calculated from the prediction error variances (PEV), so that:

$$95\%CI_{EBV} = EBV \pm 1.96\sqrt{PEV}$$

where EBV is the estimated breeding value and PEV is the prediction error variance obtained from the inverse of the coefficient matrix of the mixed model equations for random plant effects.

From the results of FA modelling, a heatmap was used to illustrate the genetic relationships and G×E interactions across the trials. First and second factor environmental loadings were plotted on the correlation scale as a uniplot. This offers greater clarity than biplots, as genotype scores and environmental loadings are plotted on separate graphs, which is important for plant breeding trials which typically deal with large numbers of genotypes and/or environments (Cullis et al. 2010). The interpretation of the environmental loadings, i.e. the directions and projections of the vectors, illustrated in a biplot (and uniplot) from a FA2 model is analogous to AMMI and other linear-bilinear models with two components (e.g. Fox et al. 1997; Yang et al. 2009). The vector length of an environment indicates the proportion of G×E variation modelled for that environment by the factors (first loading factor) and therefore its particular relationships with other vectors (environments) are the G×E variation due to disproportionality accounted for by the second loading factors (Burgueño et al. 2008). Standard errors of the environmental loadings were obtained by jackknifing; each environment was deleted in turn, data re-analysed and standard errors obtained from the re-sampled results.

6.4 Results

The mean tuber yields for plots in trials ranged from 26 t ha^{-1} (LIN-C4-05) to 70 t ha^{-1} (WAI-C5-03L), with a maximum plot yield of 139 t ha^{-1} (LIN-C4-02) (Fig. 6-2). Lincoln (LIN) trials were irrigated, which probably resulted in a greater opportunity for genotypes to better express their yield potential, as their mean yields were consistently high (56 kg ha^{-1} to 70 kg ha^{-1}) for the years 2001 to 2004. Phenotypic standard deviations for these four Lincoln trials ranged from 16 to 19 compared with a range of mostly 10 to 14 for other trials, with some exceptions falling outside this range. The low mean yield (26 t ha^{-1}) and high coefficient of variation (59) for LIN-C4-05 were probably due to water logging that was reported for a period of the growing season for this trial (plot yield range of 0 t ha^{-1} to 79 t ha^{-1}). A larger variation in mean yield (across years) was observed for Pukekohe (PK) trials compared with LIN, ranging from 28 to 50 t ha^{-1} , with coefficients of variation ranging from 23 (PK-C5-03) to 52 (PK-C4-02). For each location that trialled both C4 and C5 stages (PK, MW, LIN), observed mean yields (and usually their standard deviations) were reasonably similar within each year-location (but not across years within locations). From trial selection stages C4 to C5, observed mean yields increased and variation decreased for trials grown in the same year-location combination.

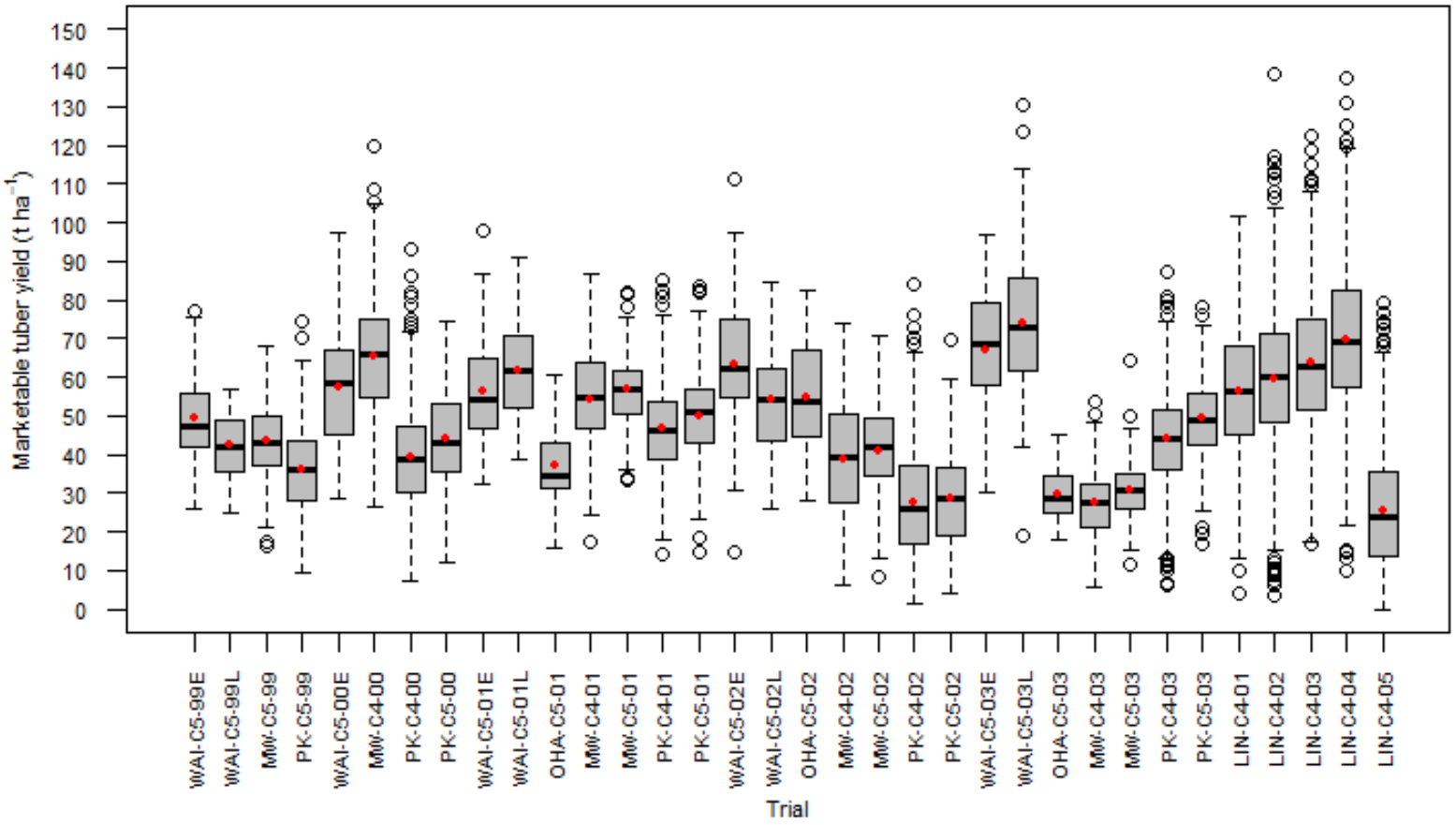

Figure 6–2 Box plots of (marketable) potato tuber yield in tonnes per hectare from 34 trials.WAI, MW, PK, OHA, LIN trial prefixes refer to Waikato, Manawatu, Pukekohe, Ohakune and Lincoln locations respectively. Trials are grouped by island (North, South) and then by year and location. Mean yields are indicated by the filled circles

Based on AIC as the model selection criterion, the FA2 variance structure for G×E provided the best fit to the data compared with the DIAG, CORH and FA1 models (Table 6-2). The total percentage variance accounted for by FA2 was a satisfactory 75%, an increase of 11% from the FA1 variance model. Starting values from the results of FA2 were used for the attempted fit of FA3 but convergence failed. Convergence also failed for the unstructured (US) model. The FA2 model was a reasonable fit for many of the trials and first latent variables were all positive (Table 10-4, Appendix V). Model fit, however, was particularly poor for trials PK-C4-01 and PK-C5-02. Model fit was also compromised, albeit to a lesser degree, for trials PK-C4-02, MW-C5-99, MW-C5-03, WAI-C5-99L, WAI-C5-01E, WAI-C5-02E, WAI-C5-03E, OHA-C5-03 and LIN-C4-05, as shown by the reasonably poor percentage of variance that was accounted for by the FA2 model (Table 10-4, Appendix V).

Table 6–2 Summary of variance models, number of variance parameters and goodness-of-fit for the multi-environment trial (MET) analysis of potato yield data

	No. variance parameters				
†G_0 structure	G_0	Total	‡AIC	-2Log-L	%variance accounted for
DIAG	34	115	1016	67185	-
CORH	35	116	114	66281	-
FA1	68	149	75	66176	64
FA2	101	182	0	66033	75

†DIAG: heterogeneous trial variances and zero covariances between pairs of trials; CORH: heterogeneous trial variances and a single correlation between pairs of trials; FA*k*: factor analytic of order *k* (1 or 2).
‡AIC expressed as the difference from the best fitting model.

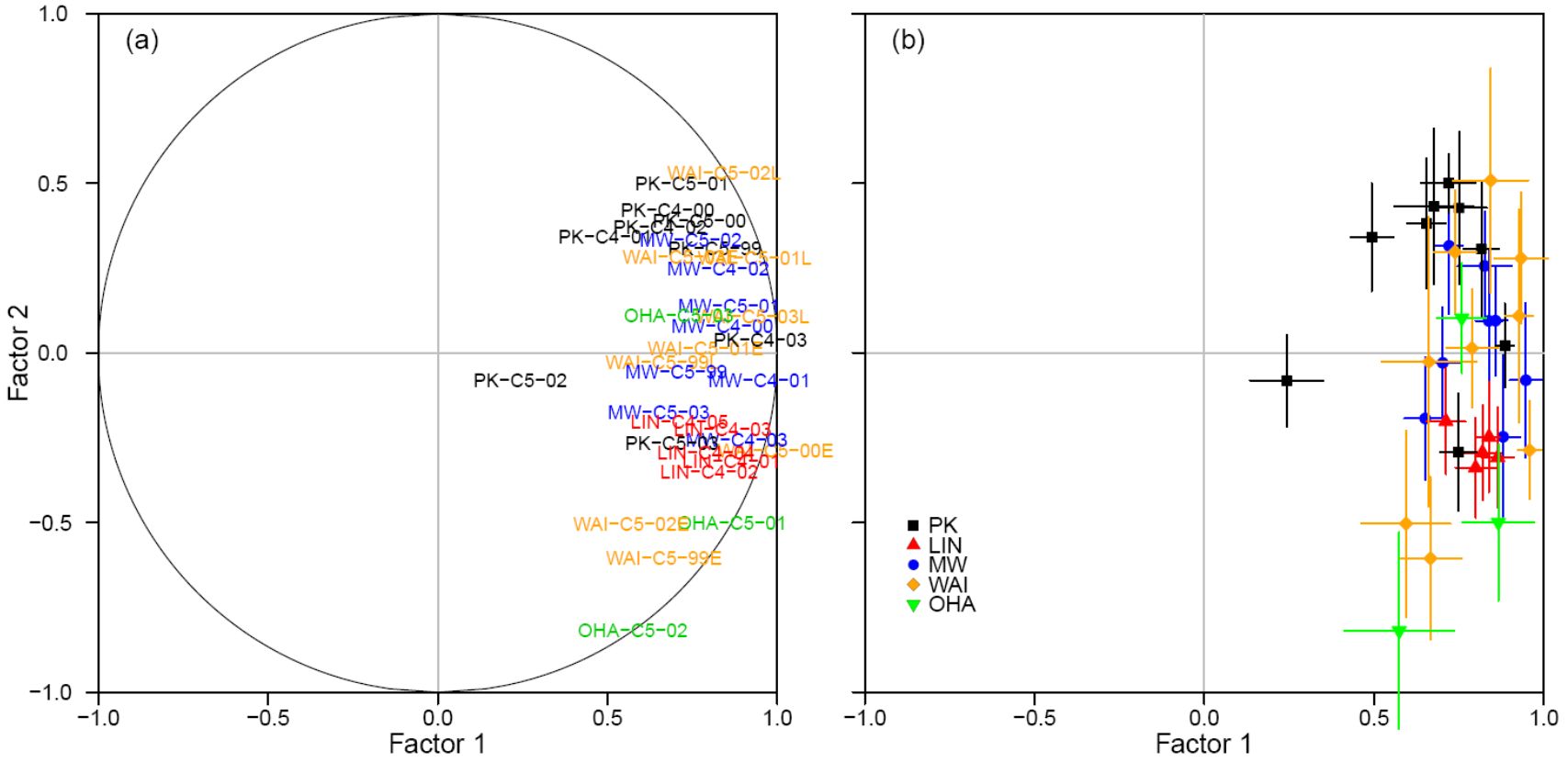

Figure 6–3 Environment uniplot of the genetic effect for potato tuber yield. Factor 1 and Factor 2 represent the rotated environment loadings on a correlation scale: (a) trials and; (b) trials with trial locations represented as symbols, and standards errors for the environmental loadings

The relationship between pairs of environments can be represented by the cosine of the angle between two environment vectors drawn from the origin (Fig. 6-3). Vectors at right angles or less represent environments for which the approximate correlation of G×E effects ranges from zero to positive. Vectors that are greater than at right angles to one another (or less than 270 degrees) approximate a negative correlation of G×E effects between environments. Therefore, environment vectors with a similar direction, as plotted from the origin, approximate a high and positive correlation, and those with opposite directions approximate a high and negative correlation (Fox et al. 1997).

The span of vectors in Figure 6.3, between WAI-C5-02L and OHA-C5-02 and therefore encompassing all trials (not drawn), is subtended by an angle of approximately 90 degrees, therefore indicating a pairwise correlation between these two trials (and a minimum correlation obtained from all pairwise combinations) of about zero. As a general rule, the uniplot (Fig. 6-3) illustrated that trials were more likely to be clustered by location rather by year. Most trials at the two main PFR research sites (PK and LIN) were delineated from one another and each tended to group together. LIN trials in particular were grouped together closely and had negative second latent variables. PK trials were mostly grouped together (with the exception of three trials) and in contrast to those of LIN trials, second latent variables were mostly positive. For trials PK-C5-02 and PK-C4-01, yields were poor and the model accounted for only 7% and 36% of their total variation respectively (Table 10-4, Appendix V). Large specific variances were also found for PK-C5-02 and (to a lesser extent) for PK-C4-01. These results indicated that the interpretation of pairwise genetic correlations inferred from the uniplot involving these two trials may be unreliable (Cullis et al. 2010). MW trials were not so tightly clustered together but fell between the two main PK and LIN clusters, which may have reflected the geographic (latitudinal) location of these three on-station trials (Fig. 6-1). For off-station trials (WAI, OHA), the clustering of locations was not so easy to discern compared with that of on-station trials (PK. MW, LIN), because of greater variation in second latent variables for WAI and OHA compared with LIN and PK. The three OHA trials, for instance, were widely dispersed with a large range of positive to negative latent variables. There may also have been a contrast between WAI (L) (positive) and WAI (E) (negative).

The genetic correlation matrix is illustrated by the heatmap shown in Figure 6-4. Genetic correlations ranged from zero to 0.97, with two groups of trials displaying particularly strong correlations. This pattern is also reflected, to a large degree, by the uniplot (and

dendrogram of the dissimilarity matrix shown in Figure 10-5, Appendix V). The reduced correlations between PK and LIN in most years are visually represented, as are the low correlations between both LIN and PK and some OHA and WAI trials. The heatmap also illustrates low correlations between PK-C5-02 (and PK-04-1) and all (or most) other trials, which is a further indication that the uniplot may not be a reliable means to infer pairwise genetic correlations involving these particular trials.

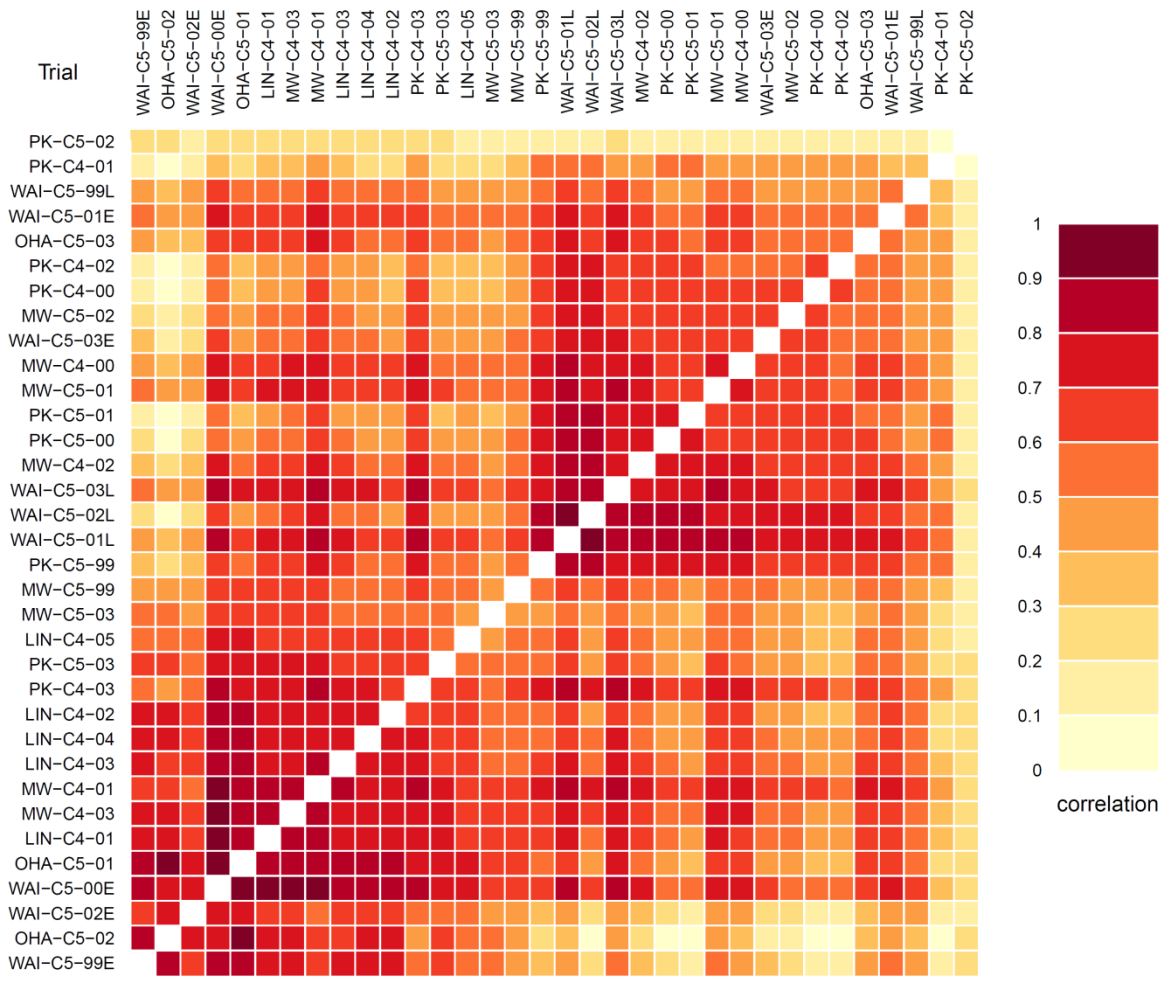

Figure 6–4 Heatmap of the genetic correlation estimates of potato tuber yield

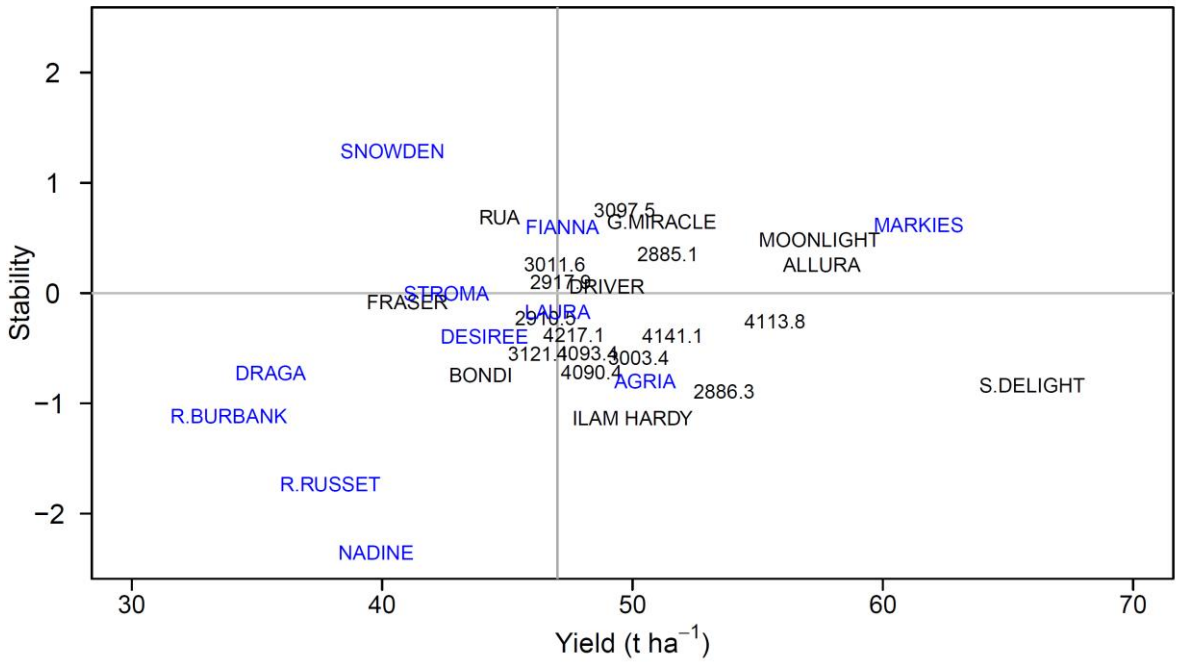

Figure 6–5 Predicted yields and stability measures (t ha^{-1}) for imported cultivars (blue) and New Zealand-bred cultivars and advanced clones (black). Standard errors of stability estimates are not shown for sake of clarity, but range from +/- 0.1 to 0.5. Confidence limits for yield are shown in Figure 6-6. The vertical line represents the New Zealand mean seasonal production yield of 46 t ha^{-1} between 2001 and 2010 (FAOSTAT 2013)

The predicted genotype values and static stability measures for a number of advanced clones and cultivars are shown in Figure 6-5. High yielding advanced clones that have been developed by the PFR breeding programme can be identified (Figs. 6-5 and 6-6), e.g. 'Moonlight', 'Allura', 'Summer Delight', but these selections did not necessarily demonstrate greater stability than recently imported cultivars. There is evidence to suggest there has been genetic improvement for potato yield in New Zealand previous to 1970, based on the comparison of more recently developed varieties with old cultivars that are still widely grown (Fig. 6-6). In 2011, the cultivars included in this analysis made up approximately 75 to 80% of the seed tuber production area (with ~33 cultivars accounting for the remainder), which indicates their importance as commercially grown cultivars in New Zealand. The pre-1970 cultivars (shown in Fig. 6-6) contributed to ~20% of the total seed tuber growing area, while three cultivars, namely 'Agria', 'Nadine' and 'Moonlight', accounted for ~40%. The international (imported) cultivars, 'Russet Burbank', 'Desiree', 'Draga', 'Nadine', 'Ranger Russet', 'Laura' and 'Agria', together contributed a large proportion of the total potato production area (~40%), but the predicted mean yield for most of these cultivars was below the mean yearly commercial yield of 46t ha^{-1} recorded from

2001 to 2010 (FAOSTAT 2013). Varieties have been developed in New Zealand that are both high yielding and acceptable in terms of cooking/processing quality. Further, many of the advanced clones developed by PFR since 2000 are above this yield threshold, suggesting that, in relation to the production cultivars presented here, the genetic selection for increased tuber yield in a multi-trait selection programme in New Zealand has been largely successful. However, the genetic advance for yield of some of the advanced clones tendered for commercial release between 2000 and 2010 has been limited and only at or just above the observed mean seasonal yield of 46t ha^{-1} reported over this period. The locally bred cultivars 'Summer Delight' and 'Allura' and the imported cultivar 'Markies' are not widely grown commercially but have been shown to be particularly high yielding under New Zealand test conditions. It should be noted that these results are based on New Zealand test conditions and assumes that there is no important genotype × management interaction, under which condition the conclusions may change.

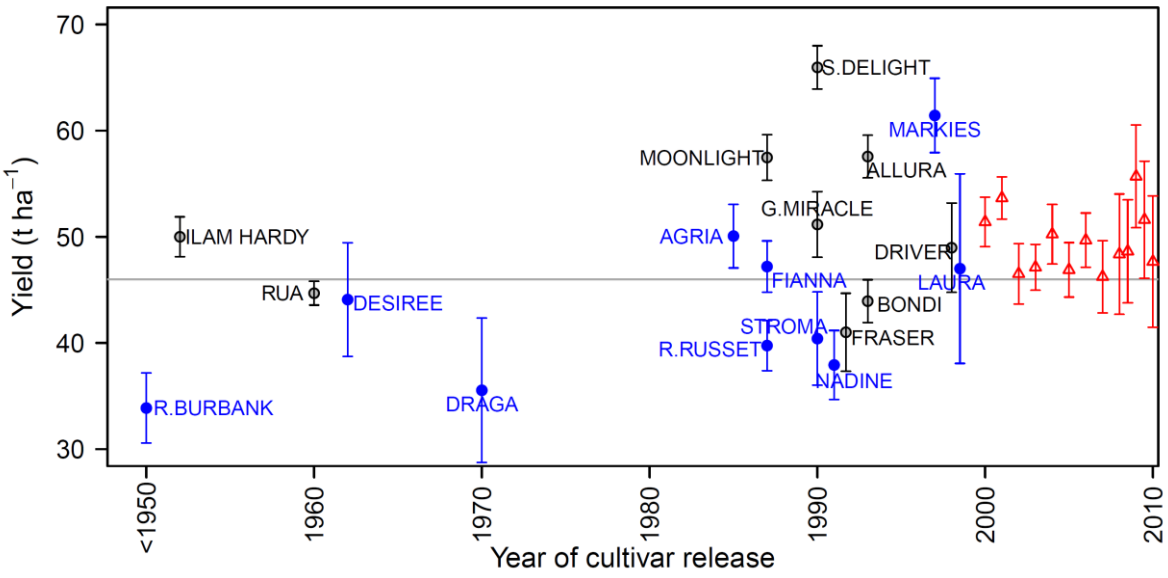

Figure 6–6 Predicted yields with 95% confidence intervals for imported potato cultivars (blue), New Zealand-bred cultivars (black) and advanced clones (△). Year of cultivar release is approximate and the horizontal line represents the New Zealand mean seasonal production yield of 46 t ha^{-1} between 2001 and 2010 (FAOSTAT 2013)

6.5 Discussion

Multi-location testing in potato evaluation trials

Yield testing in breeding programmes is highly resource-demanding in terms of land and labour requirements. The retrospective analysis of historical MET data is of interest to plant breeders as it determines the magnitude and type of G×E effects for traits, which helps to re-evaluate breeding strategies. The discrimination of trials, in terms of G×E effects, is also useful as it can provide plant breeders with information on locations with regard to differentiating genotype performance: "*Efficiency in selection necessitates rationalisation of selection locations according to similarity of selection locations in discriminating among the genotypes"* (DeLacy et al. 1996). Potato selection in the New Zealand potato breeding programme is based on genotypes that are broadly adapted to perform well over all major growing regions in New Zealand. In the present study, the main PFR trial sites, Pukekohe and Lincoln, were identified as two contrasting test locations for the evaluation of tuber yield, and that testing at these two sites (and as early as possible in the selection cycle) is likely to provide the best and most efficient opportunity to identify broadly adapted clones. Considering the three PFR research farm locations (PK, MW, LIN), there was a pattern of stratification, in general, across these 'on-station' trials (Fig. 6-3) perhaps based on latitude (Fig. 6-1), but the factors (e.g. water supply) that may have contributed to this are unknown. MW second factor environment loadings were more variable than those of LIN, effectively distributed around zero, and clustered between PK and LIN (Fig. 6-3). Genetic correlations between MW trials and other trials were positive and generally moderately-high to high (Fig. 6-4). This may suggest that MW (Manawatu) field trials are contributing little extra in terms of discriminating genotypes for tuber yield performance and broad adaptation, above and beyond that which is being achieved from comprehensive testing in Pukekohe and Lincoln; resources may be better diverted elsewhere, for example, by improving selection precision from further replication or by increasing selection intensity by evaluating more clones, at Pukekohe and Lincoln. Increasing test locations implies an increased cost of running a breeding programme if there is no net benefit in terms of, say, genetic gain or an significant increase in the probability of identifying the 'best' clones, and it incurs an opportunity cost. Dropping locations from a testing regime implies a loss of precision for the prediction of genotype values, but increasing replication to maintain precision of the genotype main effects (assuming a lack of important G×E effects) should be less costly than maintaining trials at separate locations.

Commercial potato production in New Zealand is in a temperate maritime climate with some (but not extreme) differences in rainfall, temperature and humidity. Growing regions such as Lincoln in the Canterbury region of the South Island, which are prone to long periods of water deficit in the summer months, are usually irrigated. Producers in the Pukekohe region on the other hand largely rely on rainfed production and so yields may therefore be more erratic because of greater fluctuations in seasonal water supply and season-to-season variations in precipitation. The potato crop is reported to be very sensitive to soil water condition compared with many other crops, and fluctuating water availability over the growing season can severely affect total and marketable yield (e.g. Vayda 1994 and references therein; Walworth and Carling 2002) . Off-station trials (WAI and OHA) behaved erratically in general, did not appear to group together in any predictable pattern, and genetic correlations between these trials and others were often low and effectively zero (but not consistently so). To some extent, this may have been due to poor genetic links that existed between some of these trials because of attrition of genotypes from C4 to C5, but it may also have been a reflection of off-station trials being less 'managed' than the on-station trials of PK, MW and LIN and therefore affected by year-to-year variation in, for example, water supply or other unexplained conditions outside the breeders' control. This illustrates the usefulness of regional trials grown under commercial conditions to identify well-adapted genotypes, although interaction effects are probably not repeatable and therefore less predictable (DeLacy et al. 1996). It would be informative to the breeder if there were a better understanding of the reasons for the poor correlations, and also to identify the factors that may be contributing to the stratification of the PK, MW and LIN sites (e.g. with the use of more descriptive or analytical models). With access to more extensive data, evaluation models in other crop breeding studies have aimed to characterise genotype and environment responses in relation to agronomic and abiotic factors, such as temperature and rainfall patterns (e.g. van Eeuwijk et al. 2005; Crossa et al. 2010). In Australia, Zhang et al. (2013) used such an approach to target canola breeding for specific adaptation in different environments, to either heat and drought tolerance, or high yield potential. For potato, this strategy may be particularly relevant for larger genetic improvement programmes that select specifically adapted genotypes for a much wider range of agro-climatic zones, such as the breeding schemes of the International Potato Center (CIP) in Peru.

The choice of locations for MET trials also has to take into account biotic as well as abiotic factors, which may be specific to particular regions. The Pukekohe potato crop (and most other potato growing regions in the North Island) is at risk from a disease known as

zebra chip, a relatively recent incursion (Liefting et al. 2008), which causes yield loss as well as crop rejection because of poor tuber cooking and eating quality. Therefore, variety evaluation in both the North and South Islands (where the disease, at the time of writing, is not currently endemic) is now influenced by this important biotic factor. Likewise, the Lincoln region has a lower relative humidity than the Pukekohe region during the growing season and, therefore, it is not generally affected by late blight disease (*Phytophthora infestans*), for which field resistance is an important trait in the selection criterion.

New statistical approaches have been investigated using genetic markers to better understand and predict the response of chromosome regions across different environments, e.g. QTL × environment interaction (van Eeuwijk et al. 2005) or, more recently, to predict genomic breeding values from high density markers (Heffner et al. 2009; Burgueño et al. 2012; Heslot et al. 2012) by assessing marker effect × environment interaction (Crossa 2012). A recent genomic prediction study in barley by Heslot (2013), used the prediction accuracy, rather than the genetic correlation, between environments to characterise test environments and to improve prediction of genomic breeding values in variety evaluation for the target population of environments. Studies to determine the extent and type of G×E effects found in historic tuber yield trials will help to determine the testing strategy for marker development and variety selection for potato breeding programmes in future.

Genetic improvement of potato yield and stability

Douches et al. (1996) found that the genetic yield potential of modern cultivars in the USA had not improved over those of vintage cultivars, which was attributed to a greater focus on tuber (processing) qualities and selection for early maturity. Other studies have indicated that the contribution of genetics to improve yields has been small relative to those obtained from developments in agronomic practice (Sneep and Hendriksen 1979; Walker et al. 2003), which seems to be contrary to reports for other staple crops (e.g. Duvick 2005; Mackay et al. 2011). The present study has shown that the PFR potato improvement programme has made some progress in developing advanced clones (those tendered for commercial release) and commercially released cultivars with marketable tuber yields above those found for established cultivars (Fig. 6-6) that together make up a large proportion of the current commercial crop in New Zealand (including ‘Desiree’, ‘Nadine’, ‘Agria’, ‘Russet Burbank’ and ‘Ranger Russet’). In plant variety terms, some of these cultivars are old. For example, ‘Russet Burbank’ a popular French fry cultivar, dates back to 1908 and ‘Desiree’, a popular table cultivar, dates back to 1961 (van Berloo et al. 2007).

This, to some degree, illustrates the slow adoption of newer and more productive cultivars, which is a conservatism that is reported to be characteristic of potato production systems in general (Tarn et al. 1992; Walker et al. 2003; Veilleux and De Jong 2007) and is not conducive to advancing genetic progress for tuber yield in potato production.

Point estimates for stability of a number of varieties (cultivars and advanced clones) are given in the present study (Fig. 6-5) and show that all varieties are relatively stable over the locations tested. Although estimates are likely to be often associated with large standard errors (up to +/-0.5 in the present study), particularly for genotypes tested over limited environments, such information is useful for breeders to characterise genotypes. There is no evidence to suggest that New Zealand-bred cultivars or advanced lines are any more or less stable than international cultivars bred offshore. Breeders are more likely to select promising candidates that are more consistent in performance over trials, but this is not explicitly measured in the PFR programme and stability is gauged by a genotype's variability in mean performance (relative to common standards) over many trials. From these data, it may be interesting to note that 'Russet Burbank', which is widely acknowledged to be sensitive to water stress and rapidly decreases its marketable yield because of the development of deformed tubers, was shown to be more unstable than 'Desiree' (Fig. 6-5), which is reported to be more tolerant of water stress (Vayda 1994). 'Desiree' showed near-average predicted yields and its stability possibly supports anecdotal reports that yield reliability may have, in part, contributed to its popularity. The relative yield instability of both 'Nadine' and 'Ranger Russet' compared with other tested varieties, as shown in the present study, has also been previously recognised under PFR test conditions (J. Anderson, PFR, personal communication).

With limited resources for testing genotypes, breeders largely have to ignore genotype × management interaction (G × M) effects. Anecdotally, such effects may exist, given the number of elite potato lines that meet the high expectations that breeders demand, but fail to make an impact in a commercial setting. This could, of course, be the result of numerous agronomic, economic, political and social factors, but it does illustrate the limitations of selection programmes with regard to the vast space of possible environments in which to test a restricted number of genotypes. As Messina et al. (2009) point out, the limitation with field trials is that breeders are searching a restricted set of the large space defined by all combinations of genotypes and target environments, which adds to the complication of variable management practices, expanding G and E to an even more complex G × E × M space. Breeders have to contend with G × E and largely ignore the M × E, which is left to

agronomists to deal with by identifying best management practices for elite cultivars, such as seed (tuber) storage demands, planting densities, water management and nutrient requirements. If this step is neglected, then the potential of new cultivars are often not realised if traditional management practices that apply to older established cultivars are assumed to apply also to new cultivars. Crop yield is a complex trait; to enhance rates of yield improvement and to explore the G × E × M space more effectively, a step-change in the understanding of physiological systems and processes and the development of plant simulation models has been proposed, as a means to better link the genetic variation of physiological yield determinants with their underpinning genetic systems (e.g. Hammer et al. 2006; Messina et al. 2009).

6.6 Conclusions

The evaluation of MET data from a national potato breeding programme showed that two main trial sites used for testing, Pukekohe and Lincoln (trial locations that were differentiated the most in terms of latitude), provided the best locations to distinguish the performance of varieties and to select those that were broadly adapted across target production sites in New Zealand. This study also allowed a direct comparison of the yield performance of newly developed varieties with established cultivars that are widely grown in New Zealand, to gauge genetic progress for tuber yield in the breeding programme. Selection to improve the genetic potential of tuber yield has resulted in the development and release of new cultivars that are superior to established cultivars, some of which were developed before 1970. Estimates of performance stability enable breeders to further characterise the performance of new varieties over multiple environments using MET data. Further research to relate climatic variables to genotype performance to help to interpret G×E interaction effects should go some way to improving genetic gain and targeting better deployment of specifically adapted potato cultivars. Investigating strategies to improve the selection efficiency in a potato selection programme should occur in partnership with research into molecular breeding methods, where a better understanding of G×E effects will contribute to the study and application of molecular selection methods in the genetic improvement of this important food staple.

7 General Discussion

Reviewing the established methods of field selection, and investigating strategies to improve efficiencies and reduce programme costs are necessary (but often unwelcome) features of any crop breeding programme. The analysis of historical field data provides an opportunity to explore trial heterogeneity, statistical models and trial designs that may improve the methods and precision of identifying superior genotypes for use as parents and/or cultivars. Although new molecular breeding techniques offer great promise, P. Caligari emphasised back in 1997, in an editorial of a Plant Breeding Series[1], that success will demand an integrated approach across all aspects of plant breeding, and this remark still has relevance today. The application of molecular selection methods in potato breeding is in its formative years, so far yet to deliver on the early optimism, and there has been little opportunity (or possible reason) to consider how these methods will merge with conventional selection schemes. In the meantime, this thesis has been concerned with aspects of field-based selection and developing more robust approaches to potato genetic evaluation using field data, arguably an area of research which has been largely overlooked in potato over more recent years. Some of the more salient results that have been discussed previously in Chapters 2 to 6 are re-visited, but there are also some new ideas developed and suggestions are presented for possible further work.

Genetic parameter estimation

Genetic parameter estimation is an important first step in determining a breeding strategy. Enhancing the micronutrient content of tubers has become an important selection objective at the International Potato Center (CIP) as part of a multi-faceted approach to tackle malnutrition or so-called 'hidden hunger'. This study (Chapter 2) found that for key tuber minerals iron and zinc, additive genetic effects were important, indicating that emphasis should be placed on an approach of population improvement based on recurrent selection. Heritabilities were reasonably high, indicating that individual selection will be preferred over family selection. Further, the heritabilities appeared to be inflated if the common environment of families was not taken into account. This is rarely reported in plant breeding studies, but is an interesting finding, as breeders often plant progeny families in groups, particularly in the early stages of selection (e.g. C1; Fig. 1-3). At these stages, selection is

[1]Produced by The International Centre for Advanced Mediterranean Agronomic Studies, published by Chapman & Hall

usually based on visual preference, but these results emphasise the importance of being mindful that comparisons between individuals should only be made within family groups when making selections.

There was a very favourable correlation found between iron and zinc, which is encouraging as simultaneous improvement is desirable for both. The micronutrient content of tubers reported in this study also indicates that this genetic material will provide a useful source of dietary iron and zinc, acting as a suitable base for further improvement. For example, the maximum mineral content in the G_1 progeny generation (on a fresh weight basis) was 10.4mg kg^{-1} for iron and 6.7mg kg^{-1} for zinc (data not shown). A household average consumption of 253g/ae/day (grams per adult male equivalent per day) in Peru (Rose et al. 2009) will therefore provide 56% of the estimated average requirement (EAR) of iron for children aged four to eight years and 28% for female adults aged 19 to 30, based on Dietary Reference Intakes (http://fnic.nal.usda.gov/dietary-guidance/dietary-reference-intakes/dri-tables). Similarly, consumption will provide 37% of the EAR of zinc for children aged four to eight years and 22% for female adults aged 19 to 30. However, these estimates make a number of assumptions and actual values will depend on various factors such as micronutrient bioavailability. The relationship of target micronutrients with promoters and inhibitors may affect bioavailability upon consumption (Welch and Graham 2004). Vitamin C is a known promoter of iron absorption but effectively showed a zero correlation with both iron and zinc. There are also recognised inhibitors in plant foods that suppress the absorption of iron and zinc which may compromise targets for biofortification; these include fibre, phytic acid, polyphenols and some heavy metals (Welch and Graham 2004). From human nutritional studies in common bean (*Phaseolus vulgaris*), a target crop staple for the HarvestPlus programme (CGIAR 2014), phytic acid and polyphenols have been found to inhibit the absorption of iron (Donangelo et al. 2003; Petry et al. 2012) and zinc (Donangelo et al. 2003). No such studies are reported for the consumption of potato tubers. Dietary studies and a better understanding of the genetic relationships of minerals and their known promoters and inhibitors may therefore be expedient and would be helpful to breeders, although White et al. (2009) commented that the known inhibitors phytate and oxalate are considered to be relatively low in potato tubers and therefore should not compromise the bioavailability of micronutrients.

Problems are likely to be encountered when efforts are made to enhance tuber micronutrient content along with other agronomic and tuber quality traits, using native landrace cultivars as a base population. These landraces are highly valued as food and farm

produce to the Andean people but are not adapted outside of the highland tropics. To be accepted beyond this region, undesirable characteristics such as short dormancy and deep-eyed tubers will need to be eliminated. It is hoped that the variation that exists in the diploid Andean landraces and selecting for functional levels of unreduced gametes will allow the gains obtained at the diploid level from a recurrent selection strategy to be transferred to more broadly adapted tetraploid populations.

Bayesian methods offer an alternative to REML-based approaches for the estimation of quantitative genetic parameters and breeding values. With the availability of more user-friendly software and faster computing speeds, the methods are also becoming more approachable for plant breeders. Caution is advised on their indiscriminate use, however, as the availability of plant breeding data is often limited, i.e. small breeding populations, and poor prior choice may lead to poor inferences of the posterior modes. Gelman (2006) recommended that inverse-gamma priors should not be used in hierarchical models, which are the default priors used by the MCMCglmm software (Hadfield 2010) . Alternative non-informative priors were tested on the micronutrient data in this project, following the recommendations of Gelman, but results were very similar to those found using inverse-gamma priors.

Genetic evaluation models and the use of correlated data

According to Piepho et al. (2008a), the adoption of BLUP-based selection methods by plant breeders has been slow. Similarly, it is apparent that there has been little interest in their adoption for potato genetic improvement (although this observation is based on discussions with local breeders and presumptions from literature searches). For example, a recently published text *Genetics, Genomics and Breeding of Potato*[2] gave a solitary mention to the technique, citing a single reference by Tai et al. (2009), which appears to be the first published record of BLUP methods being used in potato breeding studies. Such methods have driven animal genetic improvement in the past 30 years. The demand for BLUP in animal breeding, however, has arguably been greater (and particularly attractive) because of the widespread use of common sires through artificial insemination (across herds and across countries) and the necessity for indirect observations e.g. breeding values for milk yield in dairy sires. Further, their implementation has often been driven by large-scale data recording organisations and sire-breeding companies. Potato breeders (and possibly

[2]Edited by J.M. Bradeen and C. Kole and published in 2011 by CRC Press

breeders of many other crops also) have relied on the accumulation of data from a series of (relatively small) breeding trials but have often neglected the opportunity to improve the precision of genotype estimates by ignoring correlated information from both relatives (via the pedigree) and correlated environments (via correlation of G×E effects). Further, the availability of more user-friendly software means that there should no longer be a barrier for plant breeders in the application of more sophisticated analytical approaches. There may indeed be a certain sense of incongruity felt amongst quantitative geneticists that the growing interest from plant breeders in molecular (genotypic) selection methods and the research into genomic estimated breeding values (e.g. Heffner et al. 2009; Heslot et al. 2012) may be helping to accelerate the interest and routine adoption of phenotypic selection using BLUPs on a linear mixed modelling platform to enhance MET evaluations and field-based selection methods in plant breeding.

For ease of crop management, trials at the C2 stage in the PFR potato breeding programme (Fig. 1-3) are based on multiple trials that include a relatively small number of genotypes, i.e. 60 to 90, which are analysed independently of each other. Muller et al. (2010) commented that multiple small designs are likely to be less efficient than use of a single larger resolvable incomplete block design if selection is to be done from all entries across trials. Analysis of such designs will allow the inclusion of pedigree relationships for all candidate entries and a more straightforward incorporation of trial data from other years and locations. For pedigree-based BLUP of breeding values, it is recommended that all data that have been used in selection decisions be included in the evaluation for the estimate of less biased breeding values (Piepho and Möhring 2006). At the C2 stage, this may be appropriate for traits such as breeders' score (for the general impression of tubers) and percent marketable yield (which is highly correlated with breeders' score) but may not be necessary for tuber yield, based on the assumption that selection for yield up until this stage has been, for all intents and purposes, random (Caligari et al. 1986). It should be noted that the current series of studies used an additive relationship matrix based on disomic inheritance. The expected additive genetic covariances of both diploid and tetraploid relatives is equivalent, based on the assumptions of no past selection, double reduction or inbreeding (Lynch and Walsh 1998). These assumptions may not be reasonable. However, in a separate (as yet unpublished) analysis of potato starch data (collected on tetraploid varieties), there was very little difference in the BLUPs when a diploid relationship matrix was replaced with a tetrasomic-based relationship matrix and tested using a maximum double reduction parameter of 0.2 (P. Alspach, PFR, personal communication). The

relationship information was based on a reasonably deep pedigree (>3 generations), using the method of Kerr et al. (2012) to derive polyploid relationship matrices. Hayes et al. (2009b) replaced an expected relationship matrix built from an animal pedigree with the realised relationship matrix, constructed from dense marker genotypes, and demonstrated that the accuracy of breeding values could be notably increased.

Cowling et al. (2013) sounded a pertinent word of caution on the use of BLUPs in plant breeding programmes, citing examples of accelerated genetic gains after the introduction of BLUP evaluation methods in animal breeding schemes, but with reduced effective population sizes and unacceptably higher rates of inbreeding. By its very nature, BLUP evaluation, with its incorporation of the genetic relationship information, favours the selection of relatives as parents, which is expected given that genetic theory dictates that the best performing individuals are more likely to come from the best performing families (Simmonds 1996). As a consequence, truncation selection of BLUPs of breeding values in closed recurrent selection schemes increases the rate of inbreeding, which can be detrimental to the rate of genetic gain in the long term, leads to smaller effective population sizes, and risks greater exposure to the risk of genetic losses due to drift. In a simulated tree breeding strategy, Andersson et al. (1998) found that phenotypic selection outperformed combined index selection when constraints were imposed in order to maintain genetic diversity, and recommended restricting the number of selections from within each family. Similarly, studies in fish breeding found that selection on BLUP of breeding values resulted in unacceptable rates of inbreeding and when this was constrained, selection on phenotypic values was preferred (Sonesson et al. 2005). However, as their breeding objective included selection criteria that were measured on sibs only, it was concluded that BLUP selection was the only viable option. Previous studies (e.g. Villanueva et al. 2006) have developed methods to optimise selection on BLUP to maximise genetic gain whilst constraining the rate of inbreeding. In a potato population improvement programme, further research is required to provide breeders with guidance on the number of progenies to raise and the number of individuals to select from within each family to optimise selection and restrain inbreeding.

This study has mostly emphasised a selection objective that maximises the additive genetic response for advancing the breeding population rather than the total genetic gain for clonal deployment. Potato evaluations may also benefit from including non-additive genetic effects, which may reduce bias in breeding value estimation, and variance estimates could be exploited to by selecting favourable parental combinations (Mrode 2005). Oakey et al.

(2006) and Kelly et al. (2009) partitioned the non-additive and additive genetic effects and Oakey et al. (2007) included a dominance component in an analysis of sugarcane data. In the present study, non-additive genetic effects were not considered important when included in the evaluation of tuber micronutrients (Chapter 2) and powdery scab (Chapter 4), but to obtain satisfactory estimates of non-additive components, empirical studies tend to be based on large breeding populations (e.g. Van Tassell et al. 2000; Pante et al. 2002; Wall et al. 2005). It is also interesting to note that in autopolyploids, dominance epistasis can be passed from parent to offspring, but this is only transient and decays with further rounds of sexual propagation (Walsh 2005).

The evaluation of powdery scab (Chapter 4) using data based on a categorical scale (Fig. 10-3, Appendix III) used a linear mixed model. Despite the distribution of the phenotypic scores (Table 4-2 & Fig. 4-3), this was considered appropriate as residual quantile-quantile plots were satisfactory (Fig. 10-4, Appendix III). Linear models have, nonetheless, been shown to be reasonably robust to the non-normality of errors (Wood and Saville 2013). There are some instances where it may be considered more appropriate to analyse plant breeding data using a generalised linear mixed model, such as the analysis of dichotomous data, e.g. binary presence/absence data or categorical rating scores for disease. Evaluations using generalised linear models are less well-developed for the analysis of plant or animal breeding data, but Gianola (2007) remarked that with the availability of computing power, there is less justification nowadays to continue analysing discrete data with linear models.

There appears to be a consensus in the literature among a number of authors that spatial methods should enhance but not replace block designs altogether (Sarker et al. 2001; Piepho et al. 2008b; Müller et al. 2010). There is also evidence that spatial effects are trait-dependent (Dutkowski et al. 2006). In the potato breeding trials studied, local spatial effects did not consistently improve model fit, and the block design often appeared to deal with local heterogeneity adequately. An exception was for the analysis of p-rep trials at the Lincoln site (Chapter 3) where spatial correlations estimated from yield data were particularly high, but there was some localised water logging within the trials during the season. If local spatial trends are not strong and the trial is an incomplete design, Moder (1998; cited by Piepho et al. 2008b) reported that spatial analysis usually provided no extra advantage over classical block analysis. The recommendation from the present study concludes that spatial effects should be included on a case-by-case basis to augment the blocking features of the trial and they may be particularly helpful in the case of p-rep trials;

the extra effort involved in testing for spatial effects is trivial compared with the setting up and management of the trial itself.

Spatial correlations in the Pukekohe (replicated) early stage trials were often shown to be negative, indicating the presence of interplot competition (allocompetition or inter-genotype competition). When spatial components were important for the analysis of percent marketable yield, both negative and positive correlations were estimated (Chapter 5). It is reasonable to expect interplot competition to affect tuber yield components under resource limitations, e.g. water supply, that will vary from season to season. The initiation of tubers (tuberisation) may occur at such a rate for the potential marketable yield to be high but, under resource competition, a large proportion of tubers may not reach a critical size or mass. Pukekohe and Manawatu trials in the PFR programme are rainfed, whereas potato trials at the Lincoln site are irrigated and may therefore be less prone to water as a limiting factor to growth, with reduced year-to-year variations in water availability. Likewise, fertility trends may also affect percent marketable yield. For example, where fertility is limiting (but with no competition) spatial patchiness may result in lower percent marketable yield when tubers again fail to meet a critical size or mass because of inadequate resource capture and allocation; plot errors will not be independent but in contrast to competition, may be positive. Connolly et al. (1993) also identified competitive effects in potato yield trials. As found in the current studies, the effects were not pervasive over all trials tested. The most pertinent point from the study of Connolly et al. (1993) was that accounting for competitive effects resulted in minimal re-ranking of genotypes, but there was closer agreement with pure-stand yields, indicating that models that account for competition may be more appropriate in the advanced stages of cultivar evaluation and recommendation. Interestingly, in the same study there was no evidence of interplot competition for specific gravity, which is used to derive an estimate of the tuber dry matter content. Preliminary and unpublished studies on New Zealand trial data (and the analysis of p-rep trials in Chapter 3) have consistently found positive spatial correlations for the analysis of dry matter content. This corresponds with the findings by Connolly et al. (1993) that competitive effects of specific gravity may be absent but it may be affected by local field trends. Dry matter content is reported to be affected by the supply and soil availability of nitrogen, phosphorus and potassium (Laboski and Kelling 2007), and (Redulla et al. 2002) identified an inverse relationship between soil specific gravity and soil potassium. Soil physical and chemical properties are recognised to vary considerably across the field landscape over relatively short spatial distances (Po et al. 2010).

Autoregressive (AR1) spatial models are commonly used (and promoted) by Australian-based researchers (e.g. Gilmour et al. 1997; Stringer and Cullis 2002; Beeck et al. 2010) for crop evaluation, but the application of other nearest-neighbour adjustment methods such as linear variance models, which have been pursued by others, for example, by Piepho et al. (2008b), should not be ignored. Such methods have been shown to perform as well as or even better than AR1 models (Yang et al. 2004; Müller et al. 2010) in sugar beet, barley and pea trials. Piepho et al. (2008b) commented that simple methods of neighbour adjustment (in one direction only if blocks are orthogonal to any major field trend and consist of a single array of equally spaced plots) are often shown to be adequate and more pragmatic if routine evaluations are to be adopted.

Studies found that the trial-to-trial genetic correlations for powdery scab resistance were generally high and that a simple correlation structure (heterogeneous variances-single correlation) was adequate in modelling the G×E effects. This may simplify the routine evaluation of this trait. An interesting finding from recently published research and relevant to resistance breeding is that the variation in powdery scab pathotypes is very limited outside South America (Gau et al. 2013) . This indicates that testing for resistance in multiple locations is likely to be unnecessary and that resistance will be durable. Further, it is also likely that (outside S. America) the EBV of a genotype tested in one country will be a reliable indication of its EBV in another. Providing breeders with EBVs for powdery scab resistance will assist in enhancing genetic resistance in breeding populations. Further research is required to gain better insight into the genetic relationship between tuber infection and root galling, which will help to guide breeding strategy against the damaging *Spongospora* disease that affects both roots and tubers and compromises both tuber yield and quality.

The correlation of G×E effects found for tuber yield (Chapters 5 & 6) suggests that selection, as currently practiced in the PFR scheme, will be effective for broad adaptation of potato cultivars in New Zealand. Evaluation models using a simple correlation structure for the G×E effects were also found to be adequate for analysing early stage trials over several years when the number of test locations was limited. This is, again, more pragmatic if routine evaluations are to be adopted, particularly if the pedigree information is included, as the fitting of FA2 (and higher order models) was sometimes found to be problematic. When trials were extended to several locations at more advanced stages, the factor analytic model (FA2) was found to work well (the pedigree was ignored at these later stages). Test locations are likely to be relatively homogeneous given that both latitude and altitude

differences are small, crop water deficits that may occur in the drier eastern regions during the growing season are corrected by irrigation, and no site is routinely subjected to temperature extremes. The MET analysis of trials tested over multiple locations even suggested that there may be some duplication of sites (Fig 6-3). Consideration may therefore be given to reducing the number of tests sites, as it represents an opportunity cost where resources might be better used elsewhere. An attractive feature of assessing G×E by the approach presented is that it is relatively straightforward to perform (assuming the model will converge), as the covariance across trials is all genetic (assuming trial are appropriately randomised). In contrast, the analysis of perennial crops, such as apple or kiwifruit, has to consider environmental covariance. Hardner (2012) has explored genetic and environmental covariances in the multivariate repeated measures analysis of mango. Although the models tested were suitable for the aims of the current study, a disadvantage of the methods presented is that they are statistical rather than descriptive. Descriptive or analytical approaches aim to characterise the response of genotypes and environments in terms of abiotic and biotic factors (Fox et al. 1997), and they may also include physiological or genetical (e.g. QTL or quantitative trait loci) information. Descriptive statistical models to analyse G×E data have been reviewed by Van Eeuwijk (2005) and Crossa et al. (2010). Such approaches were used in an empirical study by Zhang et al. (2013) to characterise various environments in Australia and the response of seed yield and oil content in canola genotypes. Based on their results, phenology was found to have an influence on the performance of genotypes in contrasting climates and this information was used to develop a breeding strategy targeted at specific adaptation. Similar approaches could be followed for the evaluation of potato MET data for programmes that target specific adaptation in cultivars for a range of climates (e.g. the CIP breeding programme based in Peru).

Increasing the genetic response to selection at the early selection stages

Empirical field data were collected to measure the effectiveness of selection from partially replicated (p-rep) trials in the early stages of potato selection. Simulations also created data by sampling from normal and multivariate normal distributions to infer the genetic response by truncation selection that might be obtained from the partial replication (p-rep) of trials. As an alternative approach, the effectiveness of selection was also tested by resampling from a designed layout (using historical trial data) to maintain the integrity of the blocking structure and spatial components of the original trial. Results suggested that the use of p-rep trials at the early stages will provide an opportunity to increase the number of genotypes

that are tested at a single site, and will, for the traits that were tested at least, increase the response to selection compared with selection from fully replicated (x2) trials. P-rep trials also allow an extension of trials to multiple locations for MET testing to take account of the G×E component at an earlier stage than current practice dictates. Haynes (2012a) recommended the distribution of tubers to multiple locations in the USA at the early stages of potato selection programme to select for broadly adapted clones, but did not consider alternative p-rep or augmented designs. Even if the trial has to be replicated for the benefit of some traits (i.e. those of lower heritability), traits that do not require replication can be harvested and measured on the basis of a p-rep sampling procedure that overlays the original replicated layout. This type of approach might be useful for characteristics that are difficult or costly to measure (Smith et al. 2011).

These results have practical implications for the PFR breeding programme. Despite the fact that phenotypic selection is likely to be effective for the yield and quality traits analysed, given the phenotypic trait correlations reported in Table 3-5, the primary objective at the C1 stage is to reduce the number of genotypes to manageable numbers by eliminating those with obvious faults, such as malformed or poorly conformed tubers, or a propensity to chain tuberise. Implementing a more formal trial design and evaluation step at this stage would rely on machine harvesting and large-scale phenotyping on a small plot basis to assess the characters of interest, which would allow the screening of a greater number of genotypes. On the other hand, it would possibly incur high labour and capital costs. Phenotypic selections at C1 are grown at C2 in a replicated trial for formal evaluation at a single location, typically two replicates and 12 plants per replicate. There are enough tubers at this stage to distribute over two locations, and would allow a certain degree of replication both within and across test locations. Therefore, the main recommendation resulting from this study would be to test genotypes over two locations, namely Pukekohe and Lincoln, which are the two main trial sites in the PFR programme (Fig. 1-4), and implement a p-rep trial at the C2 stage. This would allow more genotypes to be tested at the C2 stage without increasing the number of plots required for testing. Candidates are currently tested over the two sites from the C3 or C4 stages onwards. P-rep trials would therefore also allow an earlier account of the G×E component, which can be reasonably large given that genetic correlations for marketable yield between Pukekohe and Lincoln trials were found to range from approximately 0.3 to 0.7 (Chapter 5). It may also allow more extensive testing at up to five New Zealand test locations (Fig. 6-1) at the C3 stage, a feat which is not currently achieved before the C5 to C6 stage. However, given the numbers of candidates that are still

involved at the C3 stage, this is likely to be impractical given the resource constraints of the programme.

Comments on defining a breeding objective in potato breeding programmes

Potatoes are grown in a diverse range of environments and people of different geographical regions demand various characteristics from their cultivars. Selection of multiple traits for multiple end-users may perhaps be best achieved by using customised selection indices to target end-users. For example, Bonierbale et al. (2007) identified Central Asia as a potential target region to deploy potato varieties with elevated concentrations of iron and zinc. Other traits required, as part of the overall breeding objective, were resistance to viruses and late blight (depending on the production ecology), 90- to 120-day maturity, and long dormancy with large tubers to match consumer preferences. Other regions are likely to have their own particular preferences and idiosyncrasies. Index selection requires information on the genetic control of each trait and appropriate weightings for each trait in the index. For an overall breeding objective, deriving market values to apply as economic weights in a selection index is difficult when the merit of traits is not measurable in the marketplace. This problem is encountered when developing breeding programmes for organic production systems or for developing countries where formal markets are often missing (Sölkner et al. 2008). Nutritional enhancement of potato is considered a public or society good which is not currently valued, or perhaps only partly valued or traded in the market. An alternative approach might be the use of disability-adjusted life years (DALYs) as economic weights, which were originally developed for health economic forecasting but have been adapted for the analysis of biofortified crops (Stein et al. 2005; Horton et al. 2009). This method compares the burden of the *status quo* micronutrient deficiency with the burden where micronutrient intake has increased because of the consumption of biofortified staples, and is measured in DALYs saved. Appropriate economic weighting of traits within a multiple-trait index will assist plant breeders in making informed selection decisions to maximise the response of the breeding objective.

Although breeding objectives are not reported to have been formally developed in potato breeding, an objective for the development of processing cultivars, e.g. French fry manufacture, may be more straightforward to define than an objective for the table (fresh) sector, and would put selection decisions on a more rational basis. The French fry manufacturing industry has, for many years, been dominated by the use of the 'Russet Burbank' cultivar in New Zealand, Australia and the USA, but dates back to 1914. It is

recorded to have been a chimera of the variety 'Burbank' which was, in turn, a seedling from variety 'Early Rose' and dates back to 1876 (van Berloo et al. 2007). It therefore seems apparent that selection decisions should be based on economic imperatives to better argue the case for the use of new and improved modern cultivars. The processing industry is, by and large, a vertically integrated sector, and traits in the breeding objective should be relatively easy to clarify (e.g. Fig. 1-1). Unlike farm production costs, factory manufacturing costs are, however, difficult to obtain, because of the fiercely competitive nature of the business, but conjecture could be sought by examining the energy inputs and product losses at each stage of the process. Somsen et al. (2004), for example, modelled the production process to predict losses and Orr & Graham (1983) used linear programming (as a decision support tool for potato processors) for determining the least-cost source of potatoes by relating potato productions factors, e.g. specific gravity, with energy usage and the costs of production. Tuber shape is a measured trait that also appears to have an important role is the amount of factory product loss e.g. because of unwanted peel loss and off-cuts (Fig. 1-1). It is currently measured on a categorical scale (1 to 5, or 1 to 7) with a score given for each sample (plot), but there is much intra- as well as inter-varietal variation. A more objective measure such as digital image analysis (e.g. Williams et al. 2012) may be more appropriate for this trait and such information could then be related to a prediction of factory loss and therefore go some way to providing candidates with an economic measure of their worth.

Marketable yield, as defined in this particular study (and in the PFR breeding programme in general), is a rather crude definition; a genotype's economic worth in terms of tuber yield will ultimately be dependent on the target end-use. For example, tuber size (or weight) distribution may be a factor in determining farm revenue if written into payment contracts Such a measure may therefore go some way to better describe the economic worth of genotypes, and suggests a requirement for customized yield indices depending on the specified end-user. Defining yield to maximise the selection objective for each target end-user may be relatively easy to accomplish. Actually measuring the yield trait in a selection programme may not be so straightforward given the labour and capital constraints placed on breeding programmes, as more detailed measurements, such as mean tuber weight and tuber size distribution, are likely to be required. Any customisation of yield indices and implementation of procedures, such as digital image analysis to assess tuber shape, or near infrared reflectance (e.g. Scanlon et al. 1999) to measure tuber dry matter content, suggests a requirement for automatic and high-throughput phenotyping techniques and data

collection. The accumulation of such data would assist the development of molecular breeding technologies, and also help to gain more insight into the genetic control of yield and its determinants and other agronomic and quality traits for which we have limited information. This degree of detail in accumulated data might help to provide a better understanding of the more complex G×E×M (genotype-by-environment-by-management) multi-dimensional space rather than just the G×E space (Messina et al. 2009), which, in turn, will help molecular geneticists, agronomists, crop physiologists and crop modellers, as well as plant breeders, to enhance potato productivity.

Closing remarks on the prospects for potato selection

This thesis has focused on the field-based selection procedures for potato. With the development of molecular methods (and the reduction of genotyping costs), new genotype-based selection technologies will have increasing roles to play in evaluation methods and their integration will be required for successful application in potato selection schemes. However, QTL studies based on bi-parental mapping families and genome-wide association studies (GWAS) of populations are yet to translate into the routine and widespread adoption of marker-assisted selection (MAS) methods in breeding programmes, as was once envisaged. There is an acknowledgement that most of the phenotypic variance remains unaccounted for – the so-called *missing heritability* (Marjoram et al. 2014), and for a number of years many have argued that a shift away from the overly simplistic gene-to-phenotype paradigm is required for most traits (e.g. Brady 1997) . It has been suggested that the step change required to drive further progress in crop yields and other complex traits should take a more systems-based research effort to integrate phenotypic and molecular selection methods (e.g. Hammer et al. 2006; Messina et al. 2009), and a combination of genomics and metabolomics to predict phenotypes or non-additive effects with greater precision (Gärtner et al. 2009). These methods, in themselves, inevitably possess their own complexities and, in the short term at least, plant breeders are now looking back towards a selection approach based on infinitesimal assumptions, albeit relying on dense marker genotypes, known as genomic selection (Meuwissen et al. 2001; Heffner et al. 2009) to provide genomic estimated breeding values (GEBVs). For its successful implementation in plants, there will have to be a further development of models that are more suitable for crop evaluations; methods are currently based on diploid inheritance developed for animal genetic evaluations, where the G×E component is of little interest, importance or is simply

ignored. Jonas & de Koning (2013) have reviewed the problems and prospects for genomic selection in plant breeding.

It is hoped that the approaches developed in this thesis will go some way to encourage potato breeders to review their phenotypic-based evaluation methods and selection procedures, particularly in the early stages of selection. By making better use of all their available data and exploiting the correlations that exist both within and between trials, more effective selection decisions can be made which, in turn, should increase the rate of genetic improvement and the chances of identifying superior cultivars for wide-scale deployment and cultivation.

8 General Conclusions

- To evaluate genotypes and to help guide breeding strategies, Bayesian methods offer plant breeders an alternative to REML-based procedures for the estimation of variance components, genetic parameters and empirical breeding values using an individual 'animal' model framework.

- An improved understanding of the genetic control of micronutrients iron and zinc is required to guide the breeding strategy for the biofortification of potato tubers, which is a global selection objective for a number of staple food crops. Studies indicated that additive genetic effects were important, heritabilities were moderate (but inflated if the common environment of siblings was not taken into account) and the genetic correlation between iron and zinc was strong and positive.

- An improvement strategy that employs cycles of genotypic recurrent 'individual' selection based on empirical breeding values is recommended to enhance levels of iron and zinc simultaneously in the breeding population studied.

- Partially replicated (p-rep) trials reduce the accuracy of genotype predictions but allow a greater number of candidates to be tested. Empirical and simulations studies indicated that greater rates of genetic gain could be attained from p-rep trials for a number of tuber yield and quality traits in the early stages of a New Zealand selection programme.

- Further gains might be achieved when replicating across sites (i.e. using p-rep designs within individual test sites) compared with full replication at a single site (with an equivalent number of total test plots).

- A relatively slow multiplication rate of potato tubers impedes multi-location testing of candidates. P-rep trials allow the earlier distribution of candidates across test sites thereby taking an earlier account of G × E effects compared with fully replicated trials.

- A more parsimonious genetic variance model with a simple correlation structure was often shown to be as efficient as a factor analytic model for the early stage MET (multi-environment trial) evaluation of powdery scab disease and tuber yield traits. These simple models are also easier to fit than unstructured or factor analytic models, particularly when a pedigree is included in evaluation. Simpler genetic variance structures therefore offer advantages if they are to be routinely applied to the genetic evaluation of potato breeding data.

- From the analysis of long term powdery scab resistance screening trials, the additive component of variation was important and there was no evidence of non-additive genetic effects. Narrow-sense heritabilities were moderate and the year-to-year genetic correlations were generally high. To enhance the level of resistance in the

New Zealand breeding population, greater consideration should be given to the selection of parents based on their empirical breeding values (EBV) for powdery scab resistance obtained from a MET evaluation.

- From the analysis of historic potato yield data, there was evidence of interplot competition in some years for total and marketable tuber yield.
- The blocking designed for replicated trials often appeared to deal adequately with localised heterogeneity and fitting models that include spatial effects (when they are deemed to be of importance) may make very little difference to the realised genetic gain of potato yield compared with models that ignore this information. However, spatial effects should be explored on a case-by-case basis as the effort expended is minimal compared with the effort and costs that have been involved in collecting field trial data.
- Trial locations that were differentiated the most in terms of latitude provided the best locations to distinguish the yield performance of varieties, as shown by the analysis of MET tuber yield data collected over multiple locations using a factor analytic model. Therefore, these locations are best suited to select potato genotypes that are broadly adapted across production sites in New Zealand. There was some evidence for the genetic progress, albeit somewhat limited, of tuber marketable yield in New Zealand over recent years.
- Analysis of potato breeding data using multivariate mixed models can help to guide breeding strategies, monitor genetic progress and improve resource allocation in cultivar development programmes.

9 References

Affleck I, Sullivan JA, Tarn R, Falk DE (2008) Genotype by environment interaction effect on yield and quality of potatoes. Canadian Journal of Plant Science 88: 1099-1107.

Amer P (1995) Breeding objectives - an introduction. In: 2nd European Workshop on Advanced Biometrical Methods in Animal Breeding. 12-20 June. Salzburg, pp 1-5.

Ames M, Spooner DM (2008) DNA from herbarium specimens settles a controversy about origins of the European potato. American Journal of Botany 95: 252-257.

Anderson JAD, Howard HW (1981) Effectiveness of selection in the early stages of potato breeding programs. Potato Research 24: 289-299.

Andersson EW, Spanos KA, Mullin TJ, Lindgren D (1998) Phenotypic selection can be better than selection for breeding value. Scandinavian Journal of Forest Research 13: 7-11. doi:10.1080/02827589809382956

Andre CM, Ghislain M, Bertin P, Oufir M, del Rosario Herrera M, Hoffmann L, Hausman JF, Larondelle Y, Evers D (2007) Andean potato cultivars (*Solanum tuberosum* L.) as a source of antioxidant and mineral micronutrients. Journal of Agricultural and Food Chemistry 55: 366-378. doi:10.1021/jf062740i

Annicchiarico P (2002) Genotype × environment interaction: challenges and opportunities for plant breeding and cultivar recommendations. In: Annicchiarico P (ed) FAO Plant Production and Protection Papers 174. Food and Agricultural Organization of the United Nations, Rome, pp 1-132.

Apiolaza LA, Chauhan SS, Walker JCF (2011) Genetic control of very early compression and opposite wood in *Pinus radiata* and its implications for selection. Tree Genetics & Genomes 7: 563-571. doi:10.1007/s11295-010-0356-0

Atlin GN, Baker RJ, McRae KB, Lu X (2000) Selection response in subdivided target regions. Crop Science 40: 7-13.

Atlin GN, Frey KJ (1990) Selecting oat lines for yield in low-productivity environments. Crop Science 30: 556-561.

Baldwin SJ, Genet RA, Butler RC, Jacobs JME (2008) A greenhouse assay for powdery scab (*Spongospora subterranea* f. sp *subterranea*) resistance in potato. Potato Research 51: 163-173. doi:10.1007/s11540-008-9100-7

Balzarini M (2001) Applications of mixed models in plant breeding. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CABI Publishing, Wallingford, , pp 353-363.

Bamberg JB, del Rio A (2005) Conservation of genetic resources. In: Razdan MK & A Mattoo (eds). Genetic improvement of Solanaceous crops. Science Publishers Inc., Enfield NH, pp 1-38.

Basford KE, Williams ER, Cullis BR, Gilmour A (1996) Experimental design and analysis for variety trials. In: Cooper M & GL Hammer (eds). Plant adaptation and crop improvement. CAB International, Wallingford, pp 125-138.

Beatson RA, Alspach PA (2009) The use of empirical breeding values to improve genetic progress in hops. Acta Horticulturae 848: 93-100.

Beeck CP, Cowling WA, Smith AB, Cullis BR (2010) Analysis of yield and oil from a series of canola breeding trials. Part I. Fitting factor analytic mixed models with pedigree information. Genome 53: 992-1001. doi:10.1139/g10-051

Bernardo R (1996a) Best linear unbiased prediction of maize single-cross performance. Crop Science 36: 50-56.

Bernardo R (1996b) Best linear unbiased prediction of the performance of crosses between untested maize inbreds. Crop Science 36: 872-876.

Bernardo R (2002) Breeding for quantitative traits in plants. Stemma Press, Woodbury, Minnesota.

Besag J, Kempton R (1986) Statistical analysis of field experiments using neighbouring plots. Biometrics 42: 231–251.

Bhattacharya SK, Sheo R, Dwivedi R (1985) Sources of resistance to powdery scab in potatoes. Indian Phytopathology 38: 174-175.

Birch PRJ, Bryan G, Fenton B, Gilroy EM, Hein I, Jones JT, Prashar A, Taylor MA, Torrance L, Toth IK (2012) Crops that feed the world 8: Potato: are the trends of increased global production sustainable? Food Security 4: 477-508. doi:10.1007/s12571-012-0220-1

Blasco A (2001) The Bayesian controversy in animal breeding. Journal of Animal Science 79: 2023-2046.

Bonierbale M, Amoros W, Burgos G, Salas E, Juarez H (2007) Prospects for enhancing the nutritional value of potato by plant breeding. In: African Potato Conference Proceedings. Egypt, pp 26-46.

Bos I (1983a) Optimum number of replications when testing lines or families on a fixed number of plots. Euphytica 32: 311-318.

Bos I (1983b) Some remarks on honeycomb selection. Euphytica 32: 329-335. doi:10.1007/BF00021441

Bos I, Caligari PDS (2008) Selection methods in plant breeding. 2nd edn. Springer, Dordrecht.

Bouis HE, Welch RM (2010) Biofortification: A sustainable agricultural strategy for reducing micronutrient malnutrition in the global South. Crop Science 50: S20-S32. doi:10.2135/cropsci2009.09.0531

Bradshaw J, Dale M, Mackay G (2003) Use of mid-parent values and progeny tests to increase the efficiency of potato breeding for combined processing quality and disease and pest resistance. Theoretical and Applied Genetics 107: 36-42. doi:DOI 10.1007/s00122-003-1219-y

Bradshaw J, Dale M, Mackay G (2009) Improving the yield, processing quality and disease and pest resistance of potatoes by genotypic recurrent selection. Euphytica 170: 215-227. doi:10.1007/s10681-009-9925-4

Bradshaw J, Dale M, Swan G, Todd D, Wilson R (1998) Early-generation selection between and within pair crosses in a potato (*Solanum tuberosum* subsp. *tuberosum*) breeding programme. Theoretical and Applied Genetics 97: 1331-1339.

Bradshaw J, Ramsay G (2005) Utilisation of the commonwealth potato collection in potato breeding. Euphytica 146: 9-19. doi:10.1007/s10681-005-3881-4

Bradshaw JE, Mackay GR (1994) Breeding strategies for clonally propagated potatoes. In: Bradshaw JE & GR Mackay (eds). Potato genetics. CAB International, Wallingford, pp 467-497.

Brady M (1997) Paradigm lost? Nature Biotechnology 15: 1324.

Brown CR (2008) Breeding for phytonutrient enhancement of potato. American Journal of Potato Research 85: 298-307. doi:10.1007/s12230-008-9028-0

Brown CR, Haynes KG, Moore M, Pavek MJ, Hane DC, Love SL, Novy RG, Miller JC (2010) Stability and broad-sense heritability of mineral content in potato: Iron. American Journal of Potato Research 87: 390-396. doi:10.1007/s12230-010-9145-4

Brown CR, Haynes KG, Moore M, Pavek MJ, Hane DC, Love SL, Novy RG, Miller JC (2011) Stability and broad-sense heritability of mineral content in potato: Zinc. American Journal of Potato Research 88: 238-244. doi:10.1007/s12230-011-9188-1

Brown CR, Haynes KG, Moore M, Pavek MJ, Hane DC, Love SL, Novy RG, Miller JC, Jr. (2012) Stability and broad-sense heritability of mineral content in potato: Calcium

and magnesium. American Journal of Potato Research 89: 255-261. doi:10.1007/s12230-012-9240-9

Brown J (1988) The effect of the weight of the seedling-derived tuber on subsequent clonal generations in a potato breeding programme. Annals of Applied Biology 113: 69-78.

Brown J, Caligari PDS (1986) The efficiency of seedling selection for yield and yield components in a potato breeding programme. Zeitschrift fur Pflanzenzuchtung 96: 53-62.

Brown J, Caligari PDS, Dale MFB, Swan GEL, Mackay GR (1988) The use of cross prediction methods in a practical potato breeding programme. Theoretical and Applied Genetics 76: 33-38.

Brown J, Caligari PDS, Mackay GR (1987a) The repeatability of progeny means in the early generations of a potato breeding programme. Annals of Applied Biology 110: 365-370.

Brown J, Caligari PDS, Mackay GR, Swan GEL (1984) The efficiency of seedling selection by visual preference in a potato breeding programme. Journal of Agricultural Science 103: 339-346.

Brown J, Caligari PDS, Mackay GR, Swan GEL (1987b) The efficiency of visual selection in early generations of a potato breeding programme. Annals of Applied Biology 110: 357-363.

Burgos G, Amoros W, Morote M, Stangoulis J, Bonierbale M (2007) Iron and zinc concentration of native Andean potato cultivars from a human nutrition perspective. Journal of the Science of Food and Agriculture 87: 668-675. doi:10.1002/jsfa.2765

Burgueño J, Crossa J, Cornelius PL, Trethowan R, McLaren G, Anitha K (2007) Modeling additive x environment and additive x additive x environment using genetic covariances of relatives of wheat genotypes. Crop Science 47: 311-320. doi:10.2135/cropsci2006.09.0564

Burgueño J, Crossa J, Cornelius PL, Yang R (2008) Using factor analytic models for joining environments and genotypes without crossover genotype x environment interaction. Crop Science 48: 1291-1305. doi:10.2135/cropsci2007.11.0632

Burgueño J, Crossa J, Cotes JM, San Vicente F, Das B (2011) Prediction assessment of linear mixed models for multi-environment trials. Crop Science 51: 944-954. doi:10.2135/cropsci2010.07.0403

Burgueño J, de los Campos G, Weigel K, Crossa J (2012) Genomic prediction of breeding values when modeling genotype x environment interaction using pedigree and dense molecular markers. Crop Science 52: 707-719. doi:10.2135/cropsci2011.06.0299

Butler D (2006) asreml: asreml() fits the linear mixed model. R package v.2.00.

Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual. Version 3, Queensland Department of Primary Industries & Fisheries.

Cakmak I (2008) Enrichment of cereal grains with zinc: Agronomic or genetic biofortification? Plant and Soil 302: 1-17. doi:10.1007/s11104-007-9466-3

Caligari PDS (1992) Breeding new varieties. In: Harris PM (ed) The potato crop: The scientific basis for improvement. Chapman & Hall, London, pp 334-372.

Caligari PDS, Brown J (1986) The use of univariate cross prediction methods in the breeding of a clonally reproduced crop (*Solanum tuberosum*). Heredity 57: 395-401.

Caligari PDS, Brown J, Abbott RJ (1986) Selection for yield and yield components in the early generations of a potato breeding programme. Theoretical and Applied Genetics 73: 218-222.

Caligari PDS, Brown J, Manhood CA (1985) The effect of varying the number of drills per plot and the amount of replication on the efficiency of potato yield trials. Euphytica 34: 291-296.

Cameron KC, Di HJ, Moir JL (2013) Nitrogen losses from the soil/plant system: a review. Annals of Applied Biology 162: 145-173.
Cameron ND (1997) Selection indices and prediction of genetic merit in animal breeding. CAB International, Wallingford.
CGIAR. HarvestPlus. http://www.cgiar.org/our-research/challenge-programs/harvestplus/ (accessed: 17th January 2014).
Chang YS, Milligan SB (1992) Estimating the potential of sugarcane families to produce elite genotypes using univariate cross prediction methods. Theoretical and Applied Genetics 84: 662-671.
Clarke GPY, Stefanova KT (2011) Optimal design for early-generation plant breeding trials with unreplicated or partially replicated test lines. Australian & New Zealand Journal of Statistics 53: 461-480.
Conner AJ, Barrell PJ, Baldwin SJ, Lokerse AS, Cooper PA, Erasmuson AK, Nap JP, Jacobs JME (2007) Intragenic vectors for gene transfer without foreign DNA. Euphytica 154: 341-353. doi:10.1007/s10681-006-9316-z
Connolly T, Currie I, Bradshaw J, McNicol J (1993) Inter-plot competition in yield trials of potatoes *(Solanum tuberosum* L.) with single-drill plots. Annals of Applied Biology 123: 367-377.
Coombes N (2011) DiGGer design generator under correlation and blocking. http://www.austatgen.org/files/software/downloads
Cooper M, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. Theoretical and Applied Genetics 88: 561-572.
Cotes JM, Nustez CE, Martinez R, Estrada N (2002) Analyzing genotype-by-environment interaction in potato using yield-stability index. American Journal of Potato Research 79: 211-218.
Cowling WA (2013) Sustainable plant breeding. Plant Breeding 132: 1-9.
Crossa J (2012) From genotype x environment interaction to gene x environment interaction. Current Genomics 13: 225-244.
Crossa J, Burgueño J, Cornelius PL, McLaren G, Trethowan R, Krishnamachari A (2006) Modeling genotype x environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Science 46: 1722-1733. doi:10.2135/cropsci2005.11-0427
Crossa J, Cornelius PL (2002) Linear-bilinear models for the analysis of genotype-environment interaction. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CABI Publishing, Wallingford, pp 305 - 322.
Crossa J, Fox PN, Pfeiffer WH, Rajaram S, Gauch HGJ (1991) AMMI adjustment for statistical analysis of an international wheat yield trial. Theoretical and Applied Genetics 81: 27-37.
Crossa J, Vargas M, Joshi AK (2010) Linear, bilinear, and linear-bilinear fixed and mixed models for analyzing genotype x environment interaction in plant breeding and agronomy. Canadian Journal of Plant Science 90: 561-574.
Cullis B, Gleeson AC (1991) Spatial analysis of field experiments – an extension to two dimensions. Biometrics 47: 1449–1460.
Cullis BR, Smith AB, Beeck CP, Cowling WA (2010) Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome 53: 1002-1016. doi:10.1139/g10-080
Cullis BR, Smith AB, Coombes NB (2006) On the design of early generation trials with correlated data. Journal of Agricultural Biological and Environmental Statistics 11: 381-393.

CycSoftware (2009) CycDesigN 4.0 A package for the computer generation of experimental designs. Version 4.0. CycSoftware Ltd, Hamilton, New Zealand.

DeLacy IH, Basford KE, Cooper M, Fox PN (1996) Retrospective analysis of historical data sets from multi-environment trials - theoretical development. In: Cooper M & GL Hammer (eds). Plant adaptation and crop improvement. CAB International, Wallingford, pp 243-268.

Deren CW, Alvarez J, Glaz B (1992) Use of economic criteria for selection clones in a sugarcane breeding programme. In: Proceedings of the 21st Congress of the International Society of Sugarcane Technologists. Bangkok, pp 437-449.

Dieters MJ, White TL, Littell RC, Hedge GR (1995) Application of approximate variances of variance-components and their ratios in genetic tests. Theoretical and Applied Genetics 91: 15-24.

Donangelo CM, Woodhouse LR, King SM, Toffolo G, Shames DM, Viteri FE, Cheng Z, Welch RM, King JC (2003) Iron and zinc absorption from two bean (*Phaseolus vulgaris* L.) genotypes in young women. Journal of Agricultural and Food Chemistry 51: 5137-5143. doi:10.1021/jf030151w

Douches DS, Maas D, Jastrzebski K, Chase RW (1996) Assessment of potato breeding progress in the USA over the last century. Crop Science 36: 1544-1552.

Durban M, Currie ID, Kempton RA (2001) Adjusting for fertility and competition in variety trials. Journal of Agricultural Science 136: 129-140. doi:10.1017/s0021859601008541

Dutkowski GW, Silva JCE, Gilmour AR, Lopez GA (2002) Spatial analysis methods for forest genetic trials. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 32: 2201-2214. doi:10.1139/x02-111

Dutkowski GW, Silva JCE, Gilmour AR, Wellendorf H, Aguiar A (2006) Spatial analysis enhances modelling of a wide variety of traits in forest genetic trials. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 36: 1851-1870. doi:10.1139/x06-059

Duvick DN (2005) The contribution of breeding to yield advances in maize (*Zea mays* L.). Advances in Agronomy 86: 83-145. doi:10.1016/s0065-2113(05)86002-x

Eathington S, Crosbie T, Edwards M, Reiter R, Bull J (2007) Molecular markers in a commercial breeding program. Crop Science 47: S154-S163.

Edmondson RN (2004) Past developments and future opportunities in the design and analysis of crop experiments. Journal of Agricultural Science 143: 27-33.

Egoville MJ, Sullivan JF, Kozempel MF, Jones WJ (1988) Ascorbic acid determination in processed potatoes. American Potato Journal 65: 91-97.

Evenson RE, Gollin D (2003) Crop genetic improvement in developing countries: overview and summary. In: Evenson RE & D Gollin (eds). Crop variety improvement and its effects on productivity: The impact of international agricultural research. CABI Publishing, Wallingford.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman Group Limited, Harlow, UK.

Falloon R, Curtin D, Lister R, Butler R (2005) Root function and growth of potato plants reduced by *Spongospora subterranea* infection. American Journal of Potato Research 82: 68.

Falloon RE (2008) Control of powdery scab of potato: Towards integrated disease management. American Journal of Potato Research 85: 253-260. doi:10.1007/s12230-008-9022-6

Falloon RE, Genet RA, Wallace AR, Butler RC (2003) Susceptibility of potato (*Solanum tuberosum*) cultivars to powdery scab (caused by *Spongospora subterranea* f. sp

subterranea), and relationships between tuber and root infection. Australasian Plant Pathology 32: 377-385. doi:10.1071/ap03040

Falloon RE, Merz U, Lister RA, Wallace AR, Hayes SP (2011) Morphological enumeration of resting spores in sporosori of the plant pathogen *Spongospora subterranea*. Acta Protozoologica 50: 121-132.

Falloon RE, Viljanenrollinson SLH, Coles GD, Poff JD (1995) Disease severity keys for powdery and downy mildews of pea, and powdery scab of potato. New Zealand Journal of Crop and Horticultural Science 23: 31-37.

FAO (2010) Strengthening potato value chains: Technical and policy options for developing countries, Food and Agricultural Organization of the United Nations and the Common Fund for Commodities, Rome. http://www.fao.org/docrep/013/i1710e/i1710e.pdf

FAOSTAT. http://faostat.fao.org (accessed: 28th December 2013).

Federer WT (1956) Augmented (or hoonuiaku) designs. Hawaiian Planter's Record 55: 191-208.

Fox PN, Crossa J, Romagosa I (1997) Multi-environment testing and genotype x environment interaction. In: Kempton RA & PN Fox (eds). Statistical methods for plant variety evaluation. Chapman & Hall, London, pp 117-137.

Gärtner T, Steinfath M, Andorf S, Lisec J, Meyer RC, Altmann T, Willmitzer L, Selbig J (2009) Heterosis prediction by combining information on DNA- and metabolic markers. PLoS ONE 4: e5220. doi:doi:10.1371/journal.pone.0005220

Gau RD, Merz U, Falloon RE, Brunner PC (2013) Global genetics and invasion history of the potato powdery scab pathogen, *Spongospora subterranea* f.sp *subterranea*. PLoS ONE 8: e67944. doi:10.1371/journal.pone.0067944

Gauch HG, Zobel RW (1988) Predictive and postdictive success of statistical analyses of field trials. Theoretical and Applied Genetics 76: 1-10.

Gauch HG, Zobel RW (1996) Optimal replication in selection experiments. Crop Science 36: 838-843.

Gelman A (2006) Prior distributions for variance parameters in hierarchical models. Bayesian Analysis 1: 515-533.

Genet RA, Falloon RE, Braam WF, Wallace AR, Jacobs JME, Baldwin SJ (2004) Resistance to powdery scab (*Spongospora subterranea*) in potatoes - A key component of integrated disease management. In: Nichols MA (ed) Proceedings of the First International Symposium on Root and Tuber Crops: Food Down Under, Palmerston North, New Zealand, pp 57-62.

Gianola D (2007) Inferences from mixed models in quantitative genetics. In: Balding DJ, MJ Bishop & C Cannings (eds). Handbook of statistical genetics. John Wiley & Sons, Chichester, England, pp 678-717.

Gilmour AR, Cullis BR, Verbyla AP, Gleeson AC (1997) Accounting for natural and extraneous variation in the analysis of field experiments. Journal of Agricultural Biological and Environmental Statistics 2: 269-293.

Gilmour AR, Gogel B, Cullis BR, Thompson R (2006) ASReml, User Guide. Release 2.0, VSN International Ltd, Hemel Hempstead, UK.

Gleeson AC (1997) Spatial analysis. In: Kempton RA & PN Fox (eds). Statistical methods for plant variety evaluation. Chapman & Hall, London, pp 69-85.

Gonçalves-Vidigal MC, Mora F, Bignotto TS, Munhoz REF, de Souza LD (2008) Heritability of quantitative traits in segregating common bean families using a Bayesian approach. Euphytica 164: 551-560. doi:10.1007/s10681-008-9758-6

Gonçalves E, Aubyn AS, Martins A (2007) Mixed spatial models for data analysis of yield on large grapevine selection field trials. Theoretical and Applied Genetics 115: 653-663. doi:10.1007/s00122-007-0596-z

Gopal J (2006) Considerations for successful breeding. In: Gopal J & SMP Khurana (eds). Handbook of potato production, improvement and postharvest management. CRC Press, New York.

Gopal J, Gaur PC, Rana MS (1992) Early generation selection for agronomic characters in a potato breeding program. Theoretical and Applied Genetics 84: 709-713.

Gopal J, Gaur PC, Rana MS (1994) Heritability, and intra-generation and inter-generation associations between tuber yield and its components in potato (*Solanum tuberosum* L.). Plant Breeding 112: 80-83.

Graham R, Senadhira D, Beebe S, Iglesias C, Monasterio I (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. Field Crops Research 60: 57-80. doi:10.1016/s0378-4290(98)00133-6

Gregorio GB (2002) Progress in breeding for trace minerals in staple crops. Journal of Nutrition 132: 500S-502S.

Groen AF (2003) Breeding objectives and selection strategies for layer production. In: Muir WM & SE Aggrey (eds). Poultry genetics, breeding and biotechnology. CABI Publishing, Wallingford, pp 101-112.

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. Journal of Statistical Software 33: 1-22.

Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D (2006) Models for navigating biological complexity in breeding improved crop plants. Trends in Plant Science 11: 587-593. doi:10.1016/j.tplants.2006.10.006

Hardner CM, Bally ISE, Wright CL (2012) Prediction of breeding values for average fruit weight in mango using a multivariate individual mixed model. Euphytica 186: 463-477. doi:10.1007/s10681-012-0639-7

Harrison JG, Searle RJ, Williams NA (1997) Powdery scab disease of potato - A review. Plant Pathology 46: 1-25. doi:10.1046/j.1365-3059.1997.d01-214.x

Hawkes JG (1990) The potato: evolution, biodiversity and genetic resources. Belhaven Press, London.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009a) Genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 92: 433-443. doi:10.3168/jds.2008-1646

Hayes BJ, Visscher PM, Goddard ME (2009b) Increased accuracy of artificial selection by using the realized relationship matrix. Genetics Research 91: 47-60. doi:10.1017/s0016672308009981

Hayes RJ, Thill CA (2003) Genetic gain from early generation selection for cold chipping genotypes in potato. Plant Breeding 122: 158-163. doi:10.1046/j.1439-0523.2003.00776.x

Haynes KG, Gergela DM, Hutchinson CM, Yencho GC, Clough ME, Henninger MR, Halseth DE, Sandsted E, Porter GA, Ocaya PC (2012a) Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. Euphytica 186: 573-583. doi:10.1007/s10681-012-0685-1

Haynes KG, Yencho GC, Clough ME, Henninger MR, Sterrett SB (2012b) Genetic variation for potato tuber micronutrient content and implications for biofortification of potatoes to reduce micronutrient malnutrition. American Journal of Potato Research 89: 192-198. doi:10.1007/s12230-012-9242-7

Heffner EL, Sorrells ME, Jannink JL (2009) Genomic selection for crop improvement. Crop Science 49: 1-12. doi:10.2135/cropsci2008.08.0512

Henderson CR (1976) A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. Biometrics 32: 69-83.
Heslot N, Jannink J-L, Sorrells ME (2013) Using genomic prediction to characterize environments and optimize prediction accuracy in applied breeding data. Crop Science 53: 921-933. doi:10.2135/cropsci2012.07.0420
Heslot N, Yang H-P, Sorrells ME, Jannink J-L (2012) Genomic selection in plant breeding: A comparison of models. Crop Science 52: 146-160. doi:10.2135/cropsci2011.06.0297
Hijmans R (2003) The effect of climate change on global potato production. American Journal of Potato Research 80: 271-280.
Hijmans RJ (2001) Global distribution of the potato crop. American Journal of Potato Research 78: 403-412.
Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. PLoS Genetics 4: e1000008. doi:10.1371/journal.pgen.1000008
Hils U, Pieterse L (2005) World catalogue of potato varieties. Agrimedia GmbH, Germany.
Hirschi KD (2009) Nutrient biofortification of food crops. Annual Review of Nutrition 29: 401-421. doi:10.1146/annurev-nutr-080508-141143
Hobhouse H (1987) Seeds of change: Five plants that transformed mankind. Harper and Row, London.
Horton S, Alderman H, Rivera JA (2009) Hunger and malnutrition. In: Lomborg B (ed) Global crises, global solutions. Cambridge University Press, New York.
Hospital F (2007) Challenges for effective marker-assisted selection in plants. In: 3rd International Conference of Quantitative Genetics. Aug 18-24. Hangzhou, P.R.China, pp 303-310.
IPCC (2007) Climate Change 2007: Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (AR4 Synthesis Report), Geneva, Switzerland. http://www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html
Jacobsen E, Schouten HJ (2009) Cisgenesis: an important sub-invention for traditional plant breeding companies. Euphytica 170: 235-247. doi:10.1007/s10681-009-0037-y
Jannink JL, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. (Special Issue: Plant functional genomics.). Briefings in Functional Genomics & Proteomics 9: 166-177.
Jonas E, de Koning D-J (2013) Does genomic selection have a future in plant breeding? Trends in Biotechnology 31: 497-504. doi:10.1016/j.tibtech.2013.06
Kang MS (2002a) Crop improvement: Challenges in the 21st century. Food Products Press, New York.
Kang MS (2002b) Genotype-environment interaction: Progress and prospects. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding. CABI Publishing, Wallingford, pp 221–244.
Kang MS, Gauch HG (1996) Genotype-by-environment interaction. CRC Press, London.
Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits. Stanley Thornes, Cheltenham, UK.
Kehel Z, Habash DZ, Gezan SA, Welham SJ, Nachit MM (2010) Estimation of spatial trend and automatic model selection in augmented designs. Agronomy Journal 102: 1542-1552. doi:10.2134/agronj2010.0175
Kelly AM, Cullis BR, Gilmour AR, Eccleston JA, Thompson R (2009) Estimation in a multiplicative mixed model involving a genetic relationship matrix. Genetics Selection Evolution 41 doi:10.1186/1297-9686-41-33

Kelly AM, Smith AB, Eccleston JA, Cullis BR (2007) The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Science 47: 1063-1070. doi:10.2135/cropsci2006.08.0540

Kempton RA (1984) The design and analysis of unreplicated field trials. Vortrage fur Pflanzenzuchtung: 219-242.

Kempton RA, Gleeson AC (1997) Unreplicated trials. In: Kempton RA & PN Fox (eds). Statistical methods for plant variety evaluation. Chapman & Hall, London., pp 86-100.

Kerr R, Dutkowski G, Li L, McRae T, Novy R, Schneider B, Tier B (2009) Integrated genetic analysis for potato improvement. In: Proceedings of the 14th Australasian Plant Breeding Conference. Cairns, Australia.

Kerr RJ, Li L, Tier B, Dutkowski GW, McRae TA (2012) Use of the numerator relationship matrix in genetic analysis of autopolyploid species. Theoretical and Applied Genetics 124: 1271-1282.

Khush GS (2008) Biofortification of crops for reducing malnutrition. Proceedings of the Indian National Science Academy 74: 21-25.

Kumar S, Volz RK, Alspach PA, Bus VGM (2010) Development of a recurrent apple breeding programme in New Zealand: A synthesis of results, and a proposed revised breeding strategy. Euphytica 173: 207-222.

Kutman UB, Yildiz B, Cakmak I (2011) Effect of nitrogen on uptake, remobilization and partitioning of zinc and iron throughout the development of durum wheat. Plant and Soil 342: 149-164. doi:10.1007/s11104-010-0679-5

Laboski CAM, Kelling KA (2007) Influence of fertilizer management and soil fertility on tuber specific gravity: A review. American Journal of Potato Research 84: 283–290.

Liefting LW, Perez-Egusquiza ZC, Clover GRG, Anderson JAD (2008) A new '*Candidatus Liberibacter*' species in *Solanum tuberosum* in New Zealand. Plant Disease 92: 1474. doi:10.1094/pdis-92-10-1474a

Lister R, Falloon R, Curtin D, Butler R (2004) *Spongospora subterranea* reduces host (*Solanum tuberosum*) growth. In: Proceedings of the 3rd Australasian Soilborne Diseases Symposium. South Australian Research and Development Institute, Adelaide, pp 135-136.

Lorenz AJ, Chao SM, Asoro FG, Heffner EL, Hayashi T, Iwata H, Smith KP, Sorrells ME, Jannink JL (2011) Genomic selection in plant breeding: knowledge and prospects. In: Sparks DL (ed) Advances in Agronomy, Vol 110. Elsevier Academic Press Inc, San Diego, pp 77-123.

Love SL, Werner BK, Pavek JJ (1997) Selection for individual traits in the early generations of a potato breeding program dedicated to producing cultivars with tubers having long shape and russet skin. American Potato Journal 74: 199-213.

Lulai EC, Orr PH (1979) Influence of potato specific gravity on yield and oil content of chips. American Potato Journal 56: 379-390.

Lunn DJ, Thomas A, Best N, Spiegelhalter D (2000) WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10: 325-337. doi:10.1023/a:1008929526011

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates Inc., Sunderland.

Mackay G (2007) Propagation by traditional breeding methods. Vol. 1. Potato In: Razdan MK & A Mattoo (eds). Genetic improvement of Solanaceous crops. Enfield Science Publishers, NH, pp 65-81.

Mackay I, Horwell A, Garner J, White J, McKee J, Philpott H (2011) Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and

environmental factors to trends and variability in yield over time. Theoretical and Applied Genetics 122: 225-238. doi:10.1007/s00122-010-1438-y

Marjoram P, Zubair A, Nuzhdin SV (2014) Post-GWAS: where next? More samples, more SNPs or more biology. Heredity 112: 79-88.

Maunder HB (2005) The history of potato seed certification in New Zealand, 1927-2000, The New Zealand Vegetable and Potato Growers' Federation, Wellington, New Zealand.

McNeill WH (1999) How the potato changed the world's history. Social Research 66: 67-83.

Mendoza HA (1989) Population breeding as a tool for germplasm enhancement. American Potato Journal 66: 639-653.

Merz U (2008) Powdery scab of potato - occurrence, life cycle and epidemiology. American Journal of Potato Research 85: 241-246. doi:10.1007/s12230-008-9019-1

Merz U, Falloon RE (2009) Review: powdery scab of potato - increased knowledge of pathogen biology and disease epidemiology for effective disease management. Potato Research 52: 17-37. doi:10.1007/s11540-008-9105-2

Merz U, Lees AK, Sullivan L, Schwarzel R, Hebeisen T, Kirk HG, Bouchek-Mechiche K, Hofferbert HR (2012) Powdery scab resistance in *Solanum tuberosum*: an assessment of cultivar x environment effect. Plant Pathology 61: 29-36. doi:10.1111/j.1365-3059.2011.02489.x

Messina C, Hammer G, Dong Z, Podlich D, Cooper M (2009) Modelling crop improvement in a G x E x M framework via gene-trait-phenotype relationships. In: Sadras V & D Calderini (eds). Crop physiology: Applications for genetic improvement and agronomy. Academic Press, London, pp 235-265.

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819-1829.

Meyer K (2009) Factor-analytic models for genotype x environment type problems and structured covariance matrices. Genetics Selection Evolution 41 (21) doi:10.1186/1297-9686-41-21

Milbourne D, Pande B, Bryan GJ (2007) Potato. (Genome Mapping and Molecular Breeding in Plants, Volume 3). In: Kole C (ed) Pulses, sugar and tuber crops. Springer-Verlag, Berlin, pp 205-236.

Moloney C, Griffin D, Jones PW, Bryan GJ, McLean K, Bradshaw JE, Milbourne D (2010) Development of diagnostic markers for use in breeding potatoes resistant to *Globodera pallida* pathotype Pa2/3 using germplasm derived from *Solanum tuberosum ssp andigena* CPC 2802. Theoretical and Applied Genetics 120: 679-689. doi:10.1007/s00122-009-1185-0

Moreau L, Lemarie S, Charcosset A, Gallais A (2000) Economic efficiency of one cycle of marker-assisted selection. Crop Science 40: 329-337.

Mrode RA (2005) Linear models for the prediction of animal breeding values. 2nd edn. CABI Publishing, Wallingford.

Müller BU, Kleinknecht K, Möhring J, Piepho HP (2010) Comparison of spatial models for sugar beet and barley trials. Crop Science 50: 794-802. doi:10.2135/cropsci2009.03.0153

Nestel P, Bouis HE, Meenakshi JV, Pfeiffer W (2006) Biofortification of staple food crops. Journal of Nutrition 136: 1064-1067.

NIWA. http://www.niwa.co.nz/ (accessed: 28th December 2013).

Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H (2006) Joint modeling of additive and non-additive genetic line effects in single field trials. Theoretical and Applied Genetics 113: 809-819. doi:10.1007/s00122-006-0333-z

Oakey H, Verbyla AP, Cullis BR, Wei XM, Pitchford WS (2007) Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. Theoretical and Applied Genetics 114: 1319-1332. doi:10.1007/s00122-007-0515-3

Orr PH, Graham CK (1983) A generalized model for determining least-cost sources of processing potatoes. Transactions of the ASAE 26: 1875-1883.

Ortiz R, Peloquin SJ, Freyre R, Iwanaga M (1991) Efficiency of potato breeding using FDR 2n gametes for multitrait selection and progeny testing. Theoretical and Applied Genetics 82: 602-608.

Ovchinnikova A, Krylova E, Gavrilenko T, Smekalova T, Zhuk M, Knapp S, Spooner DM (2011) Taxonomy of cultivated potatoes (*Solanum* section *Petota*: Solanaceae). Botanical Journal of the Linnean Society 165: 107-155.

Palta J (1996) Recent advances in plant responses to stress: Bridging the gap between science and technology. HortScience 31: 51-57.

Pante MJR, Gjerde B, McMillan I, Misztal I (2002) Estimation of additive and dominance genetic variances for body weight at harvest in rainbow trout, *Oncorhynchus mykiss*. Aquaculture 204: 383-392. doi:10.1016/s0044-8486(01)00825-0

Panter DM, Allen FL (1995a) Using best linear unbiased predictions to enhance breeding for yield in soybean: I. Choosing parents. Crop Science 35: 397-405.

Panter DM, Allen FL (1995b) Using best linear unbiased predictions to enhance breeding for yield in soybean: II. Selection of superior crosses from a limited number of yield trials. Crop Science 35: 405-410.

Patterson HD, Thompson R (1971) Recovery of block information when block sizes are unequal. Biometrika 31: 100-109.

Petry N, Egli I, Gahutu JB, Tugirimana PL, Boy E, Hurrell R (2012) Stable iron isotope studies in Rwandese women indicate that the common bean has limited potential as a vehicle for iron biofortification. Journal of Nutrition 142: 492-497. doi:10.3945/jn.111.149286

Pfeiffer W, McClafferty B (2007a) HarvestPlus: Breeding crops for better nutrition. Crop Science 47: S88-105.

Pfeiffer WH, McClafferty B (2007b) Biofortification: Breeding micronutrient-dense crops In: Kang MS & PM Priyadarshan (eds). Breeding major food staples. Blackwell Publishing, Ames, IA, pp 61-91.

Piepho HP (1997) Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics 53: 761-767.

Piepho HP (1998) Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. Theoretical and Applied Genetics 97: 195-201.

Piepho HP, Möhring J (2006) Selection in cultivar trials – is it ignorable? Crop Science 146: 193-202.

Piepho HP, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177: 1881-1888. doi:10.1534/genetics.107.074229

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008a) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161: 209-228. doi:10.1007/s10681-007-9449-8

Piepho HP, Richter C, Williams E (2008b) Nearest neighbour adjustment and linear variance models in plant breeding trials. Biometrical Journal 50: 164-189. doi:10.1002/bimj.200710414

Piepho HP, Williams ER (2010) Linear variance models for plant breeding trials. Plant Breeding 129: 1-8. doi:10.1111/j.1439-0523.2009.01654.x

Plaisted RL, Hoopes RW (1989) The past record and future prospects for the use of exotic potato germplasm. American Potato Journal 66: 603-627.
Plummer M (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing. Vienna, Austria.
Po EA, Snapp SS, Kravchenko A (2010) Potato yield variability across the landscape. Agronomy Journal 102: 885-894.
Potatoes NZ. Industry profile. http://www.potatoesnz.co.nz/Overview/Our-Industry/Industry-profile.htm (accessed: 28th December 2013).
Qiao CG, Basford KE, DeLacy IH, Cooper M (2000) Evaluation of experimental designs and spatial analyses in wheat breeding trials. Theoretical and Applied Genetics 100: 9-16.
R Development Core Team (2012) R: a language and environment for statistical computing. Vienna, Austria. http://www.R-project.org/
Reader J (2011) Potato: A history of the propitious esculent. Yale University Press, New Haven, CT.
Redulla CA, Davenport JR, Evans RG, Hattendorf MJ, Alva AK, Boydston RA (2002) Relating potato yield and quality to field scale variability in soil characteristics. American Journal of Potato Research 79: 317-323.
Rengel Z, Batten G, Crowley D (1999) Agronomic approaches for improving the micronutrient density in edible portions of field crops. Field Crops Research 60: 27-40.
Robinson GK (1991) That BLUP is a good thing: The estimation of random effects. Statistical Science 6: 15-32.
Rose D, Burgos G, Bonierbale M, Thiele G (2009) Understanding the role of potatoes in the Peruvian diet: An approach that combines food composition with household expenditure data. Journal of Food Composition and Analysis 22: 525-532. doi:10.1016/j.jfca.2008.10.002
Ross H (1986) Potato breeding: problems and perspectives. In: Horn W, G Robbelen & P Parey (eds). Advances in plant breeding, Suppl. 13. Paul Parey Publishing, Berlin, pp 1-132.
Sands DC, Morris CE, Dratz EA, Pilgeram AL (2009) Elevating optimal human nutrition to a central goal of plant breeding and production of plant-based foods. Plant Science 177: 377-389. doi:10.1016/j.plantsci.2009.07.011
Sarker A, Singh M, Erskine W (2001) Efficiency of spatial methods in yield trials in lentil (*Lens culinaris* ssp *culinaris*). Journal of Agricultural Science 137: 427-438. doi:10.1017/s002185960100154x
Scanlon MG, Pritchard MK, Adam LR (1999) Quality evaluation of processing potatoes by near infrared reflectance. Journal of the Science of Food and Agriculture 79: 763-771.
Schultz L, Cogan NOI, McLean K, Dale MFB, Bryan GJ, Forster JW, Slater AT (2012) Evaluation and implementation of a potential diagnostic molecular marker for H1-conferred potato cyst nematode resistance in potato (*Solanum tuberosum* L.). Plant Breeding 131: 315-321. doi:10.1111/j.1439-0523.2012.01949.x
Shah FA, Falloon RE, Butler RC, Lister RA (2012) Low amounts of *Spongospora subterranea* sporosorus inoculum cause severe powdery scab, root galling and reduced water use in potato (*Solanum tuberosum*). Australasian Plant Pathology 41: 219-228. doi:10.1007/s13313-011-0110-6
Simmonds NW (1979) Principles of crop improvement. Longman, London.
Simmonds NW (1996) Family selection in plant breeding. Euphytica 90: 201-208.

Simmonds NW, Walker DIT (1986) An economic selection index for sugarcane breeding. Euphytica 35: 311-317.

Slater AT, Cogan NOI, Forster JW (2012) Cost analysis of the application of marker-assisted selection in potato breeding. Molecular Breeding 32: 299-310. doi:10.1007/s11032-013-9871-7

Smith AB, Cullis BR, Thompson R (2001) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57: 1138-1147. doi:10.1111/j.0006-341X.2001.01138.x

Smith AB, Cullis BR, Thompson R (2005) The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. Journal of Agricultural Science 143: 449-462. doi:10.1017/s0021859605005587

Smith AB, Lim P, Cullis BR (2006) The design and analysis of multi-phase plant breeding experiments. Journal of Agricultural Science 144: 393-409. doi:10.1017/s0021859606006319

Smith AB, Thompson R, Butler DG, Cullis BR (2011) The design and analysis of variety trials using mixtures of composite and individual plot samples. Journal of the Royal Statistical Society Series C-Applied Statistics 60: 437-455. doi:10.1111/j.1467-9876.2010.00755.x

Sneep J, Hendriksen AJT (1979) Plant breeding perspectives: Centennial publication of Koninklijk Kweekbedrijf en Zaadhandel D.J. van der Have, 1879-1979. Centre for Agricultural Publishing and Documentation, Wageningen.

So YS, Edwards J (2011) Predictive ability assessment of linear mixed models in multi-environment trials in corn. Crop Science 51: 542-552. doi:10.2135/cropsci2010.06.0338

Sölkner J, Grausgruber H, Okeyo AM, Ruckenbauer P, Wurzinger M (2008) Breeding objectives and the relative importance of traits in plant and animal breeding: A comparative review. Euphytica 161: 273-282. doi:10.1007/s10681-007-9507-2

Somsen D, Capelle A, Tramper J (2004) Manufacturing of par-fried French-fries. Part 3: A blueprint to predict the maximum production yield. Journal of Food Engineering 61: 209-219.

Sonesson AK, Gjerde B, Meuwissen THE (2005) Truncation selection for BLUP-EBV and phenotypic values in fish breeding schemes. Aquaculture 243: 61-68. doi:10.1016/j.aquaculture.2004.09.024

Sorensen D, Gianola D (2002) Likelihood, Bayesian, and MCMC methods in quantitative genetics. Springer, Berlin.

Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64: 583-616. doi:10.1111/1467-9868.00353

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005) A single domestication for potato based on multilocus AFLP genotyping. Proceedings of the National Academy of Sciences of the United States of America 102: 14694-14699.

Spooner DM, Núñez J, Trujillo G, M. dRH, F. G, Ghislain M (2007) Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. Proceedings of the National Academy of Sciences of the United States of America 104: 19398-19403.

Šrek P, Hejcman M, Kunzová E (2010) Multivariate analysis of relationship between potato (*Solanum tuberosum* L.) yield, amount of applied elements, their concentrations in tubers and uptake in a long-term fertilizer experiment. Field Crops Research 118: 183-193. doi:10.1016/j.fcr.2010.05.009

Stefanova KT, Buirchell B (2010) Multiplicative mixed models for genetic gain assessment in lupin breeding. Crop Science 50: 880-891.

Stein AJ (2010) Global impacts of human mineral malnutrition. Plant and Soil 335: 133-154. doi:10.1007/s11104-009-0228-2

Stein AJ, Meenakshi JV, Qaim M, Nestel P, Sachdev HPS, Bhutta ZA (2005) Analyzing the health benefits of biofortified staple crops by means of the disability-adjusted life years approach: A handbook focusing on iron, zinc and vitamin A. In: HarvestPlus Technical Monograph Series, International Food Policy Research Institute, Washington pp 32.

Stendal C, Casler MD (2006) Maximizing efficiency of recurrent phenotypic selection for neutral detergent fiber concentration in smooth bromegrass. Crop Science 46: 297-302. doi:10.2135/cropsci2005.0083

Stringer JK, Cullis BR (2002) Application of spatial analysis techniques to adjust for fertility trends and identify interplot competition in early stage sugarcane selection trials. Australian Journal of Agricultural Research 53: 911-918. doi:10.1071/ar01151

Stringer JK, Cullis BR, Thompson R (2011) Joint modeling of spatial variability and within-row interplot competition to increase the efficiency of plant improvement. Journal of Agricultural Biological and Environmental Statistics 16: 269-281. doi:10.1007/s13253-010-0051-5

Subramanian NK, White PJ, Broadley MR, Ramsay G (2011) The three-dimensional distribution of minerals in potato tubers. Annals of Botany 107: 681-691.

Tai GCC, Coleman WK (1999) Genotype x environment interaction of potato chip colour. Canadian Journal of Plant Science 79: 433-438.

Tai GCC, Murphy AM, Xiong X (2009) Investigation of long-term field experiments on response of breeding lines to common scab in a potato breeding program. Euphytica 167: 69-76.

Tai GCC, Young DA (1984) Early generation selection for important agronomic characteristics in a potato breeding population. American Potato Journal 61: 419-434. doi:10.1007/bf02852811

Talbot M (1984) Yield variability of crop varieties in the UK. Journal of Agricultural Science 102: 315-321.

Tarn TR, Tai GCC, De Jong H, Murphy AM, Seabrook JEA (1992) Breeding potatoes for long-day, temperate climates. In: Janick J (ed) Plant Breeding Reviews. John Wiley & Sons, New York, pp 217-332.

Thompson R, Cullis B, Smith A, Gilmour A (2003) A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. Australian & New Zealand Journal of Statistics 45: 445-459. doi:10.1111/1467-842x.00297

UN Secretariat (2011) Department Of Economic And Social Affairs, Population Division. World population prospects: The 2010 revision, highlights and advance tables, United Nations, New York.

van Berloo R, Hutten RCB, van Eck HJ, Visser RGF (2007) An online potato pedigree database resource. Potato Research 50: 45-57.

van de Graaf P, Wale SJ, Lees AK (2007) Factors affecting the incidence and severity of *Spongospora subterranea* infection and galling in potato roots. Plant Pathology 56: 1005-1013. doi:10.1111/j.1365-3059.2007.01686.x

van Eeuwijk FA, Malosetti M, Yin XY, Struik PC, Stam P (2005) Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. Australian Journal of Agricultural Research 56: 883-894. doi:10.1071/ar05153

Van Tassell CP, Misztal I, Varona L (2000) Method R estimates of additive genetic, dominance genetic, and permanent environmental fraction of variance for yield and health traits of Holsteins. Journal of Dairy Science 83: 1873-1877.

Van Tassell CP, Van Vleck LD (1996) Multiple-trait Gibbs sampler for animal models: Flexible programs for Bayesian and likelihood-based (co)variance component inference. Journal of Animal Science 74: 2586-2597.

Vayda ME (1994) Environmental stress and its impact on potato yield. In: Bradshaw JE & GR Mackay (eds). Potato genetics. CAB International, Wallingford, pp 239-261.

Veilleux RE, De Jong H (2007) Potato. In: Singh RJ (ed) Genetic resources, chromosome engineering, and crop improvement. CRC Press, London, pp 17-58.

Verbyla AP, Cullis BR, Kenward MG, Welham SJ (1999) The analysis of designed experiments and longitudinal data by using smoothing splines. Journal of the Royal Statistical Society Series C-Applied Statistics 48: 269-300. doi:10.1111/1467-9876.00154

Vermeer H (1990) Optimizing potato breeding. 1. The genotypic, environmental and genotype-environment coefficients of variation for tuber yield and other traits in potato (*Solanum tuberosum* L.) under different experimental conditions. Euphytica 49: 229-236.

Villanueva B, Avenano S, Woolliams JA (2006) Prediction of genetic gain from quadratic optimisation with constrained rates of inbreeding. Genetics Selection Evolution 38: 127-146. doi:10.1051/gse:2005032

Visarada KBRS, Kanti M, Aruna C, Srujana S, Saikishore N, Seetharama N (2009) Transgenic breeding: perspectives and prospects. Crop Science 49: 1555-1563. doi:10.2135/cropsci2008.10.0581

Waldmann P (2009) Easy and flexible Bayesian inference of quantitative genetic parameters. Evolution 63: 1640-1643. doi:10.1111/j.1558-5646.2009.00645.x

Waldmann P, Ericsson T (2006) Comparison of REML and Gibbs sampling estimates of multi-trait genetic parameters in Scots pine. Theoretical and Applied Genetics 112: 1441-1451. doi:10.1007/s00122-006-0246-x

Waldmann P, Hallander J, Hoti F, Sillanpaa MJ (2008) Efficient Markov chain Monte Carlo implementation of Bayesian analysis of additive and dominance genetic variances in noninbred pedigrees. Genetics 179: 1101-1112. doi:10.1534/genetics.107.084160

Walker TS, Bi YP, Li JH, Gaur PC, Grande E (2003) Potato genetic improvement in developing countries and CIP's role in varietal change In: Evenson RE & D Gollin (eds). Crop variety improvement and its effects on productivity: The impact of international agricultural research. CABI Publishing, Wallingford, pp 315-336.

Wall E, Brotherstone S, Kearney JF, Woolliams JA, Coffey MP (2005) Impact of nonadditive genetic effects in the estimation of breeding values for fertility and correlated traits. Journal of Dairy Science 88: 376-385.

Walsh B (2005) The struggle to exploit non-additive variation. Australian Journal of Agricultural Research 56: 873-881.

Walworth JL, Carling DE (2002) Tuber initiation and development in irrigated and non-irrigated potatoes. American Journal of Potato Research 79: 387-395.

Wastie RL (1991) Resistance to powdery scab of seedling progenies of *Solanum tuberosum*. Potato Research 34: 249-252.

Welch RM, Graham RD (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective. Journal of Experimental Botany 55: 353-364. doi:10.1093/jxb/erh064

White J, Zasoski R (1999) Mapping soil micronutrients. Field Crops Research 60: 11-26.

White P, Bradshaw J, Dale M, Ramsay G, Hammond J, Broadley M (2009) Relationships between yield and mineral concentrations in potato tubers. HortScience 44: 6-11.

White PJ, Broadley MR, Hammond JP, Ramsay G, Subramanian NK, Thompson J, Wright G (2012) Bio-fortification of potato tubers using foliar zinc-fertiliser. Journal of Horticultural Science & Biotechnology 87: 123-129.

White TL, Adams WT, Neale DB (2007) Forest Genetics. CABI Publishing, Wallingford.

Wilkinson GN, Eckert SR, Hancock TW, Mayo O (1983) Nearest neighbour (NN) analysis of field experiments. Journal of the Royal Statistical Society 45: 151–211.

Williams E, Piepho HP, Whitaker D (2011) Augmented p-rep designs. Biometrical Journal 53: 19-27. doi:10.1002/bimj.201000102

Williams ER (1986) A neighbor model for field experiments. Biometrika 73: 279–287.

Williams K, Munkvold J, Sorrells M (2012) Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.). Euphytica 190: 99-116. doi:10.1007/s10681-012-0783-0

Windhausen VS, Wagener S, Magorokosho C, Makumbi D, Vivek B, Piepho HP, Melchinger AE, Atlin GN (2012) Strategies to subdivide a target population of environments: Results from the CIMMYT-led maize hybrid testing programs in Africa. Crop Science 52: 2143-2152. doi:10.2135/cropsci2012.02.0125

Wood GR, Saville DJ (2013) A geometric examination of linear model assumptions. Australian and New Zealand Journal of Statistics 55: 285-303.

Xu X, Pan S, Cheng S, Zhang B, Mu D, Ni P, Zhang G, Yang S, Li R, Wang J, Orjeda G, Guzman F, Torres M, Lozano R, Ponce O, Martinez D, De la Cruz G, Chakrabarti SK, Patil VU, Skryabin KG, Kuznetsov BB, Ravin NV, Kolganova TV, Beletsky AV, Mardanov AV, Di Genova A, Bolser DM, Martin DMA, Li G, Yang Y, Kuang H, Hu Q, Xiong X, Bishop GJ, Sagredo B, Mejia N, Zagorski W, Gromadka R, Gawor J, Szczesny P, Huang S, Zhang Z, Liang C, He J, Li Y, He Y, Xu J, Zhang Y, Xie B, Du Y, Qu D, Bonierbale M, Ghislain M, del Rosario Herrera M, Giuliano G, Pietrella M, Perrotta G, Facella P, O'Brien K, Feingold SE, Barreiro LE, Massa GA, Diambra L, Whitty BR, Vaillancourt B, Lin H, Massa AN, Geoffroy M, Lundback S, DellaPenna D, Buell CR, Sharma SK, Marshall DF, Waugh R, Bryan GJ, Destefanis M, Nagy I, Milbourne D, Thomson SJ, Fiers M, Jacobs JME, Nielsen KL, Sonderkar M, Iovene M, Torres GA, Jiang J, Veilleux RE, Bachem CWB, de Boer J, Borm T, Kloosterman B, van Eck H, Datema E, Hekkert BtL, Goverse A, van Ham RCHJ, Visser RGF, Potato Genome Sequencing C (2011) Genome sequence and analysis of the tuber crop potato. Nature 475: 189-195. doi:10.1038/nature10158

Yang RC, Crossa J, Cornelius PL, Burgueño J (2009) Biplot analysis of genotype x environment interaction: Proceed with caution. Crop Science 49: 1564-1576. doi:10.2135/cropsci2008.11.0665

Yang RC, Ye TZ, Blade SF, Bandara M (2004) Efficiency of spatial analyses of field pea variety trials. Crop Science 44: 49-55.

Zhang HP, Berger JD, Milroy SP (2013) Genotype x environment interaction studies highlight the role of phenology in specific adaptation of canola (*Brassica napus*) to contrasting Mediterranean climates. Field Crops Research 144: 77-88. doi:10.1016/j.fcr.2013.01.006

Zuckerman L (1999) The potato: How the humble spud rescued the Western world. 1st edn. North Point Press, New York.

10 Appendices

Appendix I Biofortification: crossing designs and pedigree structure

For the crossing designs outlined in Chapter 2, Table 10-1 shows the structure of the G_1 population that was based on a nested mating design (North Carolina Design I or NCD I). The pedigree of the G_1 population is shown in Figure 10-1 showing the base parental generation derived from Groups Phureja, Goniocalyx and Stenotomum. Likewise, Table 10-2 shows the structure of the G_2 population based on a factorial mating design (NCD II).

Table 10–1 Structure of the G_1 (NCD I) crossing design; parents and family sizes

Female CIP identity	**Male CIP identity**				
	703287	**703421**	**703825**	**704218**	**family**
702736	34	x	x	x	303892
703280	35	x	x	x	303888
703312	36	x	x	x	303887
703317	34	x	x	x	303886
702815	x	31	x	x	303846
703291	x	36	x	x	303845
703825	x	36	x	x	303842
704393	x	27	x	x	303841
701165	x	x	35	x	303835
703168	x	x	23	x	303832
703352	x	x	34	x	303828
703421	x	x	23	x	303827
703831	x	x	29	36	303826 | 303798
700313	x	x	x	35	303806
703197	x	x	x	36	303803
704481	x	x	x	36	303797

Table 10–2 Structure of the G_2 (NCD II) crossing design; parents and family sizes

Female CIP identity	**Male CIP identity**			
SET 1	**303803.161**	**303828.201**	**303835.111**	**303888.41**
303887.101	43	53	66	61
303887.111	50	74	53	56
303887.171	22	44	45	25
303887.61	30	55	56	38
SET 2	**303803.181**	**303832.141**	**303835.191**	**303888.151**
303826.21	37	36	28	38
303826.41	47	49	40	23
303841.22	19	27	39	46
303846.11	28	42	21	38

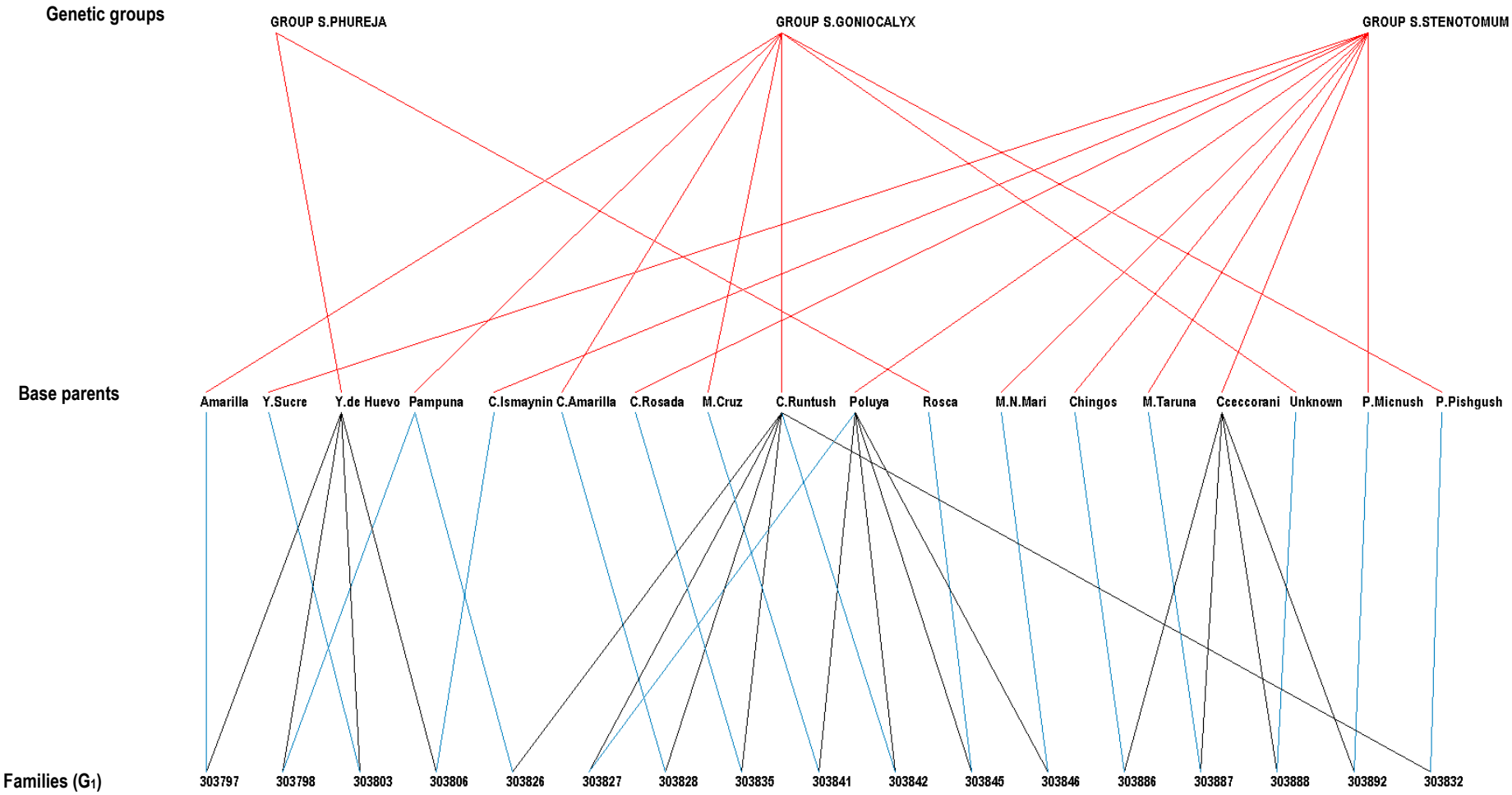

Figure 10–1 Structure of pedigree showing genetic groups, base parents and G_1 families

Appendix II P-rep trial designs

```
# Initialise Design ###########
SEEDS                          # RNG seeds [-1 -1 [-1 -1]]
238   RANDOM 1.0           # No. treatments, [F-fixed/R-random [gamma]]
24 12   24 12                 # Layout: Design r c, Replicate r c
READ SWAP                       # G-GENERATE/R-READ 1st design [S-SWAP file]

# Search Phase Settings #######
  24 12   24 12               # Swap r c, Correlation r c
  Agg                         # A-type
  0.0 20000                   # Target A-measure, Max. interchanges to test
  100.0                       # Search intensity, (low) 0 - 100 (high)
  2                           # Number of objectives
#   Objective Specifications ##
    0.8                       # Objective weight
      2                       #  No. of block factors
         24    6   1.0        #  Block r c gamma
          8   12   1.0        #  Block r c gamma
      NONE                    #  None/Row/Column/Both Linear trend code
      ID                      #  Correlation between rows: Model [par.]
      ID                      #  Correlation between columns: Model [par.]
      1.0                     #  Spatial gamma
    0.2                       # Objective weight
      1                       #  No. of block factors
          8    6   1.0        #  Block r c gamma
      NONE                    #  None/Row/Column/Both Linear trend code
      ID                      #  Correlation between rows: Model [par.]
      ID                      #  Correlation between columns: Model [par.]
      1.0                     #  Spatial gamma

# DiGGer control ##############
NEW PHASE                     # N - NEW PHASE or E - END search here

# Search Phase Settings #######
  8 6   24 12                 # Swap r c, Correlation r c
  A11                         # A-type
  0.0 100000                  # Target A-measure, Max. interchanges to test
  100.0                       # Search intensity, (low) 0 - 100 (high)
  1                           # Number of objectives
#   Objective Specifications ##
    1.0                       # Objective weight
      3                       #  No. of block factors
          1   12   1.0        #  Block r c gamma
         24    1   0.5        #  Block r c gamma
          1    1   0.1        #  Block r c gamma
      NONE                    #  None/Row/Column/Both Linear trend code
      AR  0.25                #  Correlation between rows: Model[par.]
      AR  0.25                #  Correlation between columns: Model [par.]
      1.0                     #  Spatial gamma

# DiGGer control ##############
END                           # N - NEW PHASE or E - END search here
```

Figure 10–2 Input file for the design of $S2_1$ p-rep trial (Fig.3-1) using DiGGer (Coombes 2011)

Figure 10-2 shows the input file provided for the DiGGer trial design software (Coombes 2011) for the design of the p-rep trial $C2_1$. The resulting trial is displayed in Chapter 3 (Fig. 3-1). The $C2_2$ trial was also a p-rep design but based on 21 rows and 12 columns.

Appendix III Powdery scab data and analysis

The powdery scab score card used for the 12 years of disease screening trials is illustrated in Figure 10-3. Tuber assessment for powdery scab severity with a single score assigned to each plot based upon visual assessment of all tubers was adapted for resistance screening from the scoring scheme described by Falloon, (1995). Tuber assessment was scored on an ordinal 0 to 9 scale, where 0 = no visible lesions and 9 = complete surface area covered by powdery scab lesions

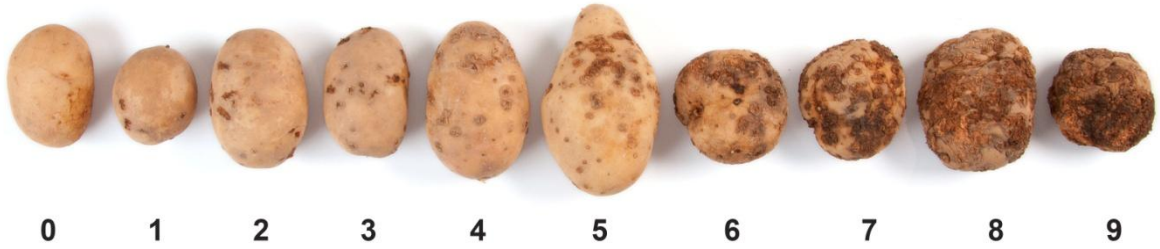

Figure 10–3 Powdery scab score card on a 0 to 9 scale where 0 = no visible symptoms and 9 = complete tuber coverage of powdery scab lesions

Residual quantile-quantile plots for the single trial analysis of powdery scab data are shown in Figure 10-4 for the 12 years of data collected between 1998 and 2010 (no available data for 2006).

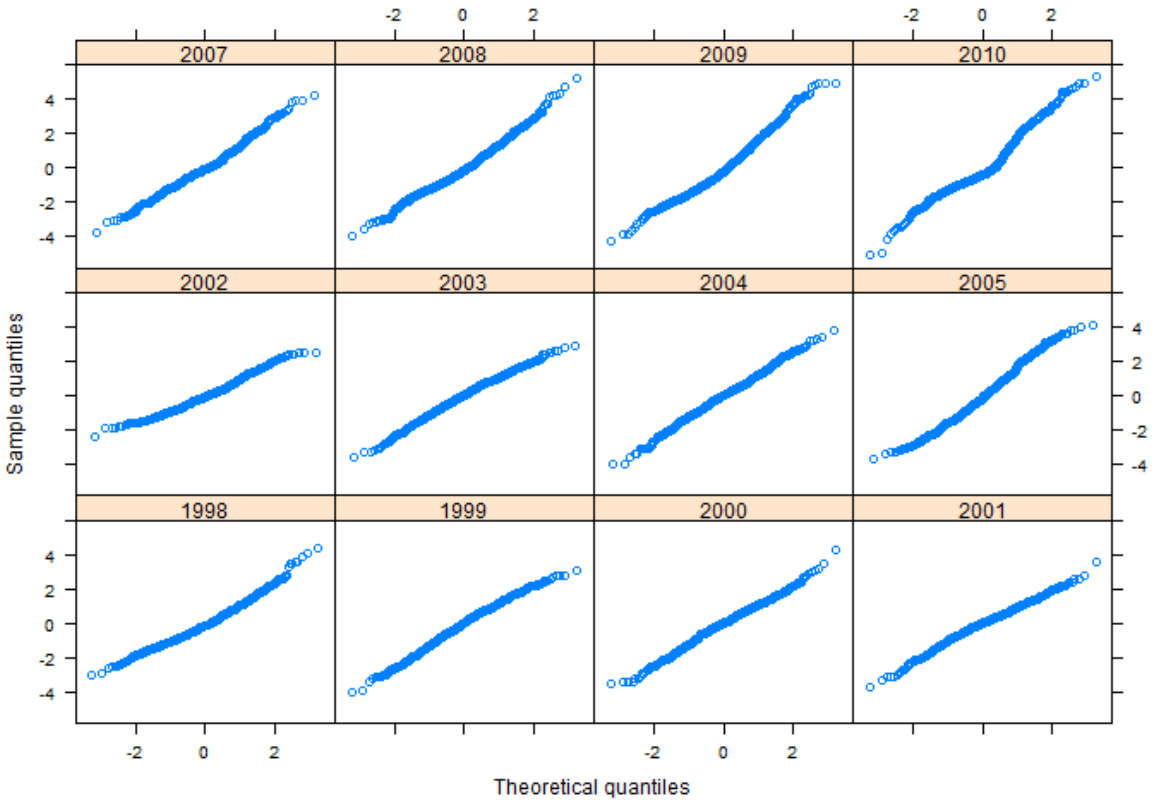

Figure 10–4 Quantile-quantile plots of residuals from single trial analysis of powdery scab data 1998 to 2005 and 2007 to 2010

Appendix IV Summary statistics of tuber yield trial data

Potato yield data for 21 years of trial data analysed in Chapter 5 is summarised in Table 10-3. Trial codes in bold showed reasonable genotype concurrence and were included in the MET analysis that included 15 trials. Distributions of potato yield data are presented as box plots in Figure 5-1.

Table 10–3 Summary of New Zealand (North Island) potato breeding trials for total yield (TTY, t ha^{-1}), marketable yield (MTY, t ha^{-1}), and commercial harvest index (CHI%, untransformed)

[a]Trial code	Year	Dimension	Genotypes	TTY			MTY			CHI		
				mea	[b]sd	[c]CV%	mean	sd	CV%	mean	sd	CV%
PK-C2-99A	1999	40 x 16	314	56	18	31	36	14	40	63	11	17
PK-C2-99B	1999	38 x 8	150	55	16	29	31	12	39	56	13	23
PK-C2-00A	2000	60 x 8	236	72	17	24	62	16	26	86	7	8
PK-C2-00B	2000	14 x 8	56	67	16	24	55	15	27	81	9	11
PK-EM-00	2000	33 x 8	86	77	14	19	66	12	19	86	7	8
PK-MN-00A	2000	20 x 24	235	52	15	29	39	13	35	73	10	14
PK-MN-00B	2000	14 x 8	56	57	14	25	43	12	28	76	9	11
PK-EM-01	2001	24 x 8	64	47	8	17	41	8	19	87	6	7
PK-MN-01A	2001	12 x 8	48	57	15	27	45	14	31	79	8	11
PK-MN-01B	2001	34 x 8	134	55	13	23	47	11	24	86	7	9
MW-MN-01	2001	12 x 8	48	61	16	25	54	14	26	90	6	7
PK-EM-02	2002	24 x 12	64	46	15	33	38	14	36	81	8	10
PK-MN-02	2002	14 x 8	56	43	18	40	34	15	44	79	9	12
MW-MN-02	2002	12 x 10	60	48	15	32	42	15	35	87	7	8
PK-EM-03	2003	30 x 8	79	71	15	21	64	14	22	90	5	5
PK-MN-03	2003	18 x 8	47	59	12	20	49	11	23	83	6	7
PK-C2-06A	2006	54 x 10	266	49	11	23	39	11	29	79	10	13
PK-C2-06B	2006	26 x 10	123	52	12	23	41	11	27	80	9	11
PK-C2-07	2007	34 x 20	324	52	9	16	44	8	19	84	7	8
PK-C2-12A	2012	56 x 7	192	64	13	20	55	13	23	86	8	9
PK-C2-12B	2012	54 x 7	187	65	14	22	56	14	24	87	8	9

[a]Trial codes in bold italics were used in the MET analysis to test different variance structures. PK or MW refers to the trial location (Pukekohe or Manuwatu) and C2 (2nd clonal stage trial), EM ('early-main') or MN ('main') describes the yield trial type. [b]sd is the standard deviation and [c]CV% is the coefficient of variation as a percentage.

Appendix V Analysis of tuber MET yield data using a factor analytic model

Table 10–4 Estimates of rotated environment loadings (first and second (scaled) latent variables, ℓ_1 and ℓ_2 respectively) and the percentage of variance accounted for (%V) by the first latent variable (ℓ_1) and the first and second latent variable ($\ell_1 + \ell_2$)

Trial	ℓ_1	ℓ_2	%V (ℓ_1)	%V (ℓ_1+ℓ_2)
WAI-C5-99E	0.67	-0.61	44	81
WAI-C5-99L	0.66	-0.03	44	44
MW-C5-99	0.70	-0.03	49	49
PK-C5-99	0.82	0.31	67	76
WAI-C5-00E	0.96	-0.29	92	100
MW-C4-00	0.84	0.09	70	71
PK-C4-00	0.68	0.43	46	65
PK-C5-00	0.75	0.43	56	75
WAI-C5-01E	0.79	0.01	62	62
WAI-C5-01L	0.93	0.28	87	95
OHA-C5-01	0.87	-0.50	75	100
MW-C4-01	0.95	-0.08	90	90
MW-C5-01	0.86	0.10	74	74
PK-C4-01	0.49	0.34	24	36
PK-C5-01	0.72	0.50	52	77
WAI-C5-02E	0.59	-0.50	35	61
WAI-C5-02L	0.84	0.51	71	97
OHA-C5-02	0.57	-0.82	33	100
MW-C4-02	0.83	0.26	68	75
MW-C5-02	0.72	0.32	52	62
PK-C4-02	0.65	0.38	43	57
PK-C5-02	0.24	-0.08	6	7
WAI-C5-03E	0.74	0.30	55	63
WAI-C5-03L	0.93	0.11	86	87
OHA-C5-03	0.76	0.10	58	59
MW-C4-03	0.88	-0.25	77	84
MW-C5-03	0.65	-0.19	42	46
PK-C4-03	0.89	0.02	78	79
PK-C5-03	0.75	-0.29	56	65
LIN-C4-01	0.86	-0.31	75	84
LIN-C4-02	0.80	-0.34	64	75
LIN-C4-03	0.84	-0.25	70	76
LIN-C4-04	0.82	-0.29	67	76
LIN-C4-05	0.71	-0.20	51	55

Table 10-4 displays the REML estimates of the (scaled) latent variables (rotated loadings) and the percentage variance accounted for in each trial (for each factor) from the MET analysis of Chapter 6. The FA2 model was a reasonable fit for many of the trials and first latent variables were all positive, but trials PK-C4-01 and PK-C5-02 performed particularly poorly. Model fit was also compromised, albeit to a lesser degree, for trials PK-C4-02, MW-C5-99, MW-C5-03, WAI-C5-99L, WAI-C5-01E, WAI-C5-02E, WAI-C5-03E, OHA-C5-03 and LIN-C4-05.

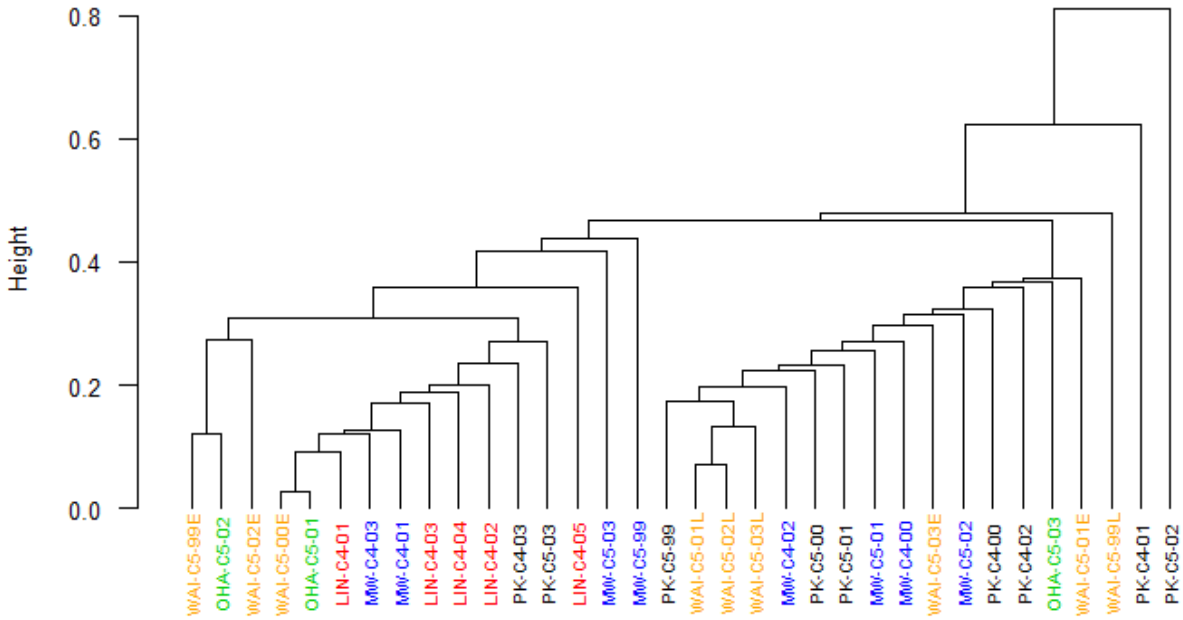

Figure 10–5 Height of dissimilarity for the genetic effects between early-stage potato selection trials for MTY (marketable yield) from a FA2 model

Figure 10-5 presents a dendrogram as a graphical alternative to the heatmap in Figure 6-4 (in the same trial order), showing the relationship between trials or 'environments'. This shows a similar trial 'clustering' pattern of two possible main groups and supports the results presented by both the heatmap (which shows the genetic correlations between trials) in Figure 6-4 and the environmental loadings in Figure 6-3. It is formed from a nested hierarchical clustering algorithm, implemented in the 'agnes' package of R (R Development Core Team 2012) and outlined by Cullis et al. (2010, p.1004). Both PK-C4-01 and PK-C5-02 merge with high combination dissimilarity which indicates that there is little correlation between these two trials and most of the others.